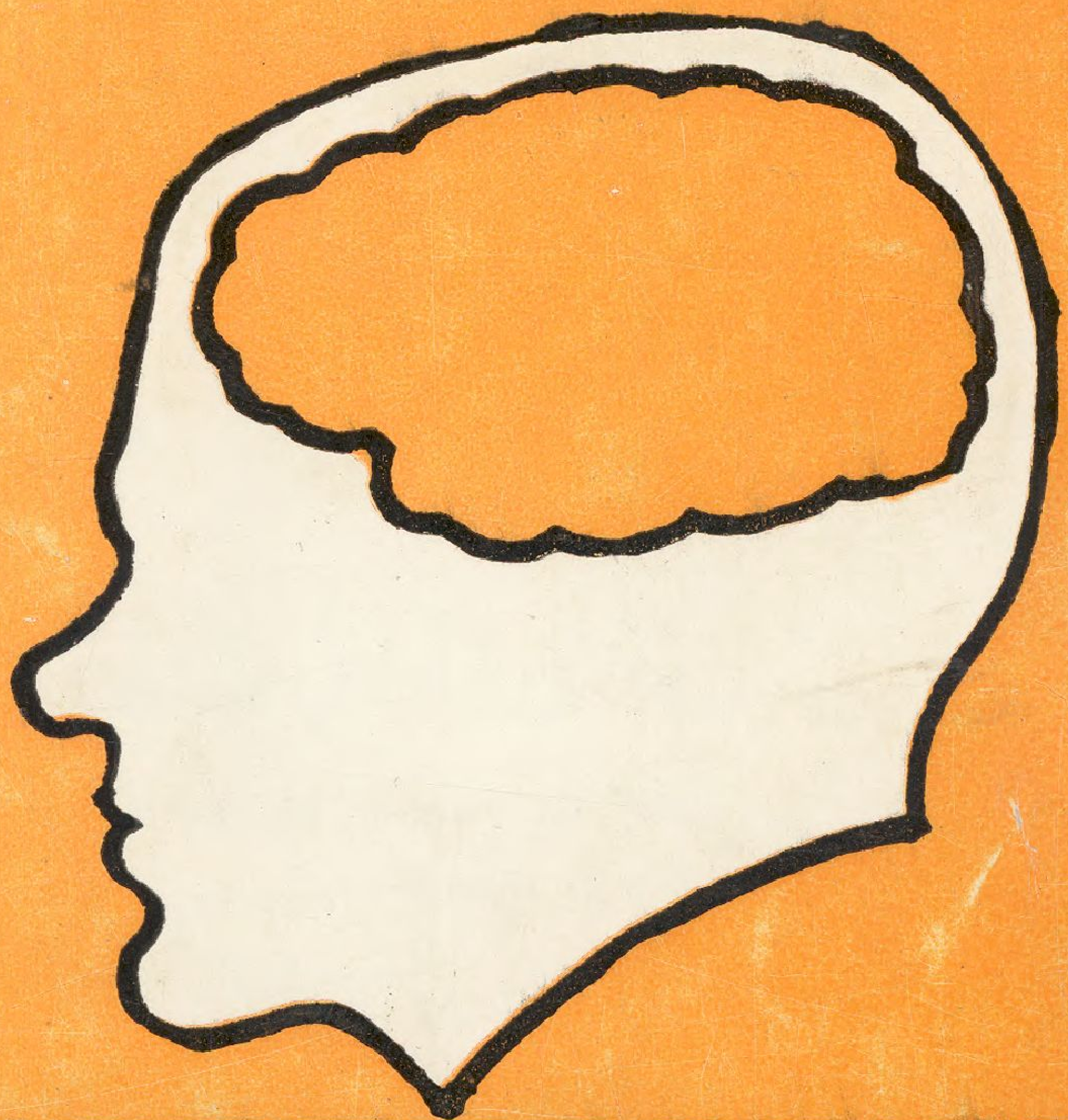
التقويم النفسي



دكتور فؤاد أبو حطب
دكتور سيد أحمد عثمان

دكتورة آمال صادق

التقويم النفسي

تأليف

دكتور فؤاد أبو حطب

أستاذ علم النفس
قسم علم النفس التعليمي
كلية التربية جامعة عين شمس

دكتور سيد أحمد عثمان

أستاذ علم النفس
قسم علم النفس التعليمي
كلية التربية جامعة عين شمس

دكتورة آمال صادق

أستاذ علم النفس ووكيلة كلية التربية
جامعة حلوان

١٩٩٣

الناشر
مكتبة الأنجلو المصرية
١٦٥ ش محمد فريد

صدر هذا الكتاب لأول مرة بعنوان (مشكلات التقويم النفسى) عام ١٩٧٠ ثم صدر بعنوان (التقويم النفسى) كما يلى :

الطبعة الأولى	١٩٧٣
الطبعة الثانية	١٩٧٦
اعادة طبع	١٩٧٩
اعادة طبع	١٩٨٣
اعادة طبع	١٩٨٤
اعادة طبع	١٩٨٥
اعادة طبع	١٩٨٦
الطبعة الثالثة	١٩٨٧

رقم الايداع ٢٢١١ / ١٩٩٣

بسم الله الرحمن الرحيم

تقديم الطبعة الثالثة

عندما صدر كتابنا الصغير المركز (مشكلات فى التقويم النفسى) فى مطلع عام ١٩٧٠ لم نكن نتوقع له أن يتطور وينمو على النحو الذى صار اليه • لقد استقبل هذا الكتاب يومئذ من دارسى علم النفس فى مصر والوطن العربى. استقبالا ابهجنا وشجعنا على أن نولى ميدان التقويم النفسى مزيدا من الاهتمام • ولم تنقض سنوات ثلاث حتى نفد هذا الكتاب الصغير وصدرت فى عام ١٩٧٣ الطبعة الأولى من كتاب (التقويم النفسى) فى صورته الأكثر شمولا واتساعا وتقدما من الكتاب السابق • ولم تمر سنوات ثلاث أخرى الا وصدرت فى عام ١٩٧٦ الطبعة الثانية من هذا الكتاب تضيف الى الطبعة الأولى وتصحح بعض ما جاء فيها •

وشاءت ظروفنا أن تمضى السنوات العشر الماضية دون أن تصدر طبعة جديدة من هذا الكتاب ، وكل ما نشر فى أعوام ١٩٧٩ ، ١٩٨٢ ، ١٩٨٤ ، ١٩٨٥ ، ١٩٨٦ لم يكن سوى اعادة اصدار للطبعة الثانية ، بعضــــه ظهر بعلمنا وعن ناشرنا ، وبعضه الآخر كان تصويرا وتزويرا للكتـــاب قام به ناشرون أخرون • وظاهرة تصوير الكتب وتزويرها أصبحت جزءا خطيرا من ازمة النشر التى يتعرض لها الكتاب المصرى فى السنوات الأخيرة ، زاد من سعارها اقبال القارىء على الكتاب الجيد • وعلى الرغم من ان هذا يتضمن فى ذاته حكما بالجودة على كتابنا الا أنه فى نفس الوقت يحمل نذر خطر اخلاقى يهدد بالشر المستطير •

كان انقطاعنا عن تجديد الكتاب طوال السنوات العشر الماضية يسبب لنا الاما مضنية • فقضايا التقويم النفسى « متعددة متجددة » كما وصفناها منذ صدر كتاب (مشكلات فى التقويم النفسى) • وقد حددنا منذ ذلك الحين مظاهر هذا التعدد والتجدد وحصرناها فى ثلاثة جوانب : أولاها طبيعـــة الظواهر النفسية موضع التقويم ، وثانيتها طبيعة أدوات التقويم النفسى ،

وثالثتها طبيعة المشكلات التى تنشأ عن استخدام هذه الأدوات فى مجالات التطبيق ·

والحق أن هذه الجوانب الثلاثة شهدت ـ منذ منتصف السبعينات ـ ولا تزال تشهد تطورات وتغيرات هامة يكاد يصل بعضها الى حد « الثورة العلمية » الكاملة ولعل أكثر مظاهر هذا التطور عمقا ما طرأ على النظرية السيكومترية العامة من اتساع لآفاقها وامتداد لحدودها جعلها تتجاوز « الاختبارات المرجعة الى المعيار » الى « الاختبارات المرجعة الى المحك » ، وعلى نحو تطلب اعادة النظر فى كثير من المفاهيم الكلاسيكية لنظرية القياس النفسى ·

أما أكثر مظاهر التطور حدة فى ميدان التقويم النفسى فهو ما يشهده مجال التطبيق من قضايا خلافية ظهرت على وجه الخصوص فى المجتمعات التى تعد الاختبارات النفسية جزءا من بنيتها المعرفية العامة ، وعلى وجه أخص بريطانيا والولايات المتحدة الأمريكية · وكان لهذا الجـــــدل الحاد تضميناته الاجتماعية والثقافية والأيديولوجية ·

يضاف الى ذلك ما يمكن أن نسميه « أزمة الاختبارات النفسية فى مصر والوطن العربى » ·وهى أزمة معقدة متعددة الجوانب تشمل فيما تشمل ندرة الابداع فى هذا المجال السيكولوجى الهام ، والاعتماد فى كثير مما ظهر من الاختبارات النفسية على النقل المباشر أو الترجمة الحرفية أو المحرفة عن اللغات الأوربية ـ وخاصة اللغة الانجليزية · ناهيك عن النقص المتزايد فى اعداد المتخصصين المتمرسين فى هذا الميدان، مع زيادة مطردة فى أعداد الهـــواة ·

ولعل أكثر ما يدعو الى الضيق حقا أن التقويم النفسى الذى يمتد عمره الزمنى فى مصر الى أكثر من خمسين عاما لم يتجاوز بعد الأروقة الأكاديمية الضيقة · وهو وضع نجد ما يشبهه فى جميع الأقطار العربية · ولعل هذا يدفعنا الى طرح تساؤل ربما يكون موضوعا لبحث علمى جاد : هل أسباب ذلك فينا نحن السيكولوجيين الذين عجزنا عن أن نصنع قنوات اتصال جيدة مع ثقافتنا ؟ أم أنها فى مجتمعاتنا التى لا تزال تتجنب العلم سبيلا الى حل مشكلاتها ؟ أم هى فينا وفى مجتمعاتنا معا ·

كان هذا كله وغيره وراء تفكيرنا فى اصدار هذه الطبعة الجديدة من كتاب التقويم النفسى والتى تطلبت اعادة النظر فى بنيته ومحتواه جميعا ٠ كما تطلبت أن تشـــاركنا بجهودها فيها أ٠د٠ آمال أحمـــد مختار صادق فأعيد تنظيمه فى ستة أبواب ٠ تناول الباب الأول الأسس العامة للتقـــويم النفسى وجاء فى فصلين أحدهما حول ما هو التقويم النفسى ، وثانيهما يقدم للقارىء المبتدىء مدخلا مبسـطا للاحصاء النفسى ٠ وقد تناولنا فى الفصل الأول معظم القضايا الخلافية المعاصرة حول الاختبارات النفسية ، أما الفصل الثانى فهو جديد تماما وقد أضفناه الى هذه الطبعة لحاجة القارىء اليه ٠

أما الباب الثانى فيعرض لمبادىء القياس النفسى ، وجـــاء فى أربعة فصول تناولت على التوالى موضوعات : فن اعطاء الاختبارات ، والثبات ، والصدق ، وتصحيح الاختبارات وتفسير الدرجات ٠ وقد تناول هذا الباب معظم الجديد فى النظرية السيكومترية العامة والذى شـــهدته طوال العقد المنقضى ٠

وقد خصص الباب الثالث « لاختبارات الذكاء العام » وينقســـم الى فصلين احدهما حول اختبارات الذكاء الفردية وعلى وجه أخص اختبارات ستانفرد ـ بينيه ووكسلر ، وثانيهما حـــول اختبارات الذكاء الجماعية مع تناول مفصل لهذه الاختبارات بنوعيها والبحوث المصرية والعـــربية التى اجريت حولها ٠

أما الباب الرابع فحول بطاريات الاستعدادات وينقسم أيضا الى فصلين احدهما حول بطاريات الاستعدادات المتعددة ، وثانيهما حـــول اختبارات القدرات المتخصصة وقد عرضنا هذه الاختبارات بالتفصيل الكافى مع تناول الجهود المصرية والعربية التى بذلت فى بناء بعض هذه الاختبارات أو فى تقنيتها ٠

وأفردنا فى هذه الطبعة الباب الخامس لاختبارات التحصيل ، وجاء هذا الباب فى فصلين تناول أولهما بناء الاختبار التحصيلى ، أما ثانيهما فحول الاختبارات المرجعة الى المحك ، وهو فصل جديد يتناول هذا الاتجاه الهام فى القياس التربوى ٠

أما الباب السادس فيتناول مقاييس الشخصية ، ويتألف من أربعة فصول تناولت على التوالى : وسائل التقرير الذاتى ، استخدام الأحكام والملاحظات المنظمة ، استخدام الأداء والأساليب الاسقاطية ، وقياس الميول ٠

وختمنا كل باب من أبواب الكتاب الستة بمراجعه ، ثم ختمنا الكتاب كله بقائمة مركزة تساعد المشتغل بالتقويم النفسى على تقويم أدواته التى يستعملها ٠

ويسعى الكتاب من خلال ذلك كله الى أن يكسب القارىء والدارس والباحث فى علم النفس بصيرة أعمق بميدانه وبذاته وبأدواته من خلال تناول الحقائق العلمية وعرض القضايا الخلافية ، ومن خلال الاشارة الى بعض المزالق والمخاطر التى يمكن أن يقع فيها ، ومن خلال التنبيه الى بعض الأخطاء التى وقع فيها بالفعل غيره ، ولعلنا بذلك نكون قد ابلغنا اللهم فاشهد ٠

القاهرة :
٤ سبتمبر ١٩٨٦ م
غرة المحرم ١٤٠٧ هـ

المؤلفون

فهـرس الكتاب

البـــاب الأول

الأسس العـــامة

الفصـــل الأول

ما هو التقويم النفسى ؟

معنى كلمة تقـــويم :

من الكلمات التى شغلت الباحثين والعامة على حد سواء فى السنوات الأخيرة كلمة « تقويم » ، وهل لا تزال تدل على المعنى المراد منها ؟ أم الأصح أن نستخدم كلمة « تقييم » والتى أصبحت أخف على الألسنة وأوسع فى الانتشار ؟

ولعل مصدر الخلاف هو الجذر الثلاثى للكلمة وهو « قوم » ومنه تصاغ كلمة « قيمة » تبعا للقاعدة الصرفية فى اللغة العربية التى ترى أن « الواو اذا وقعتساكنة بعد حرفمكسور قلبتياء لتناسب الكسرة» (١١٢:١٨٠،١٢ب، ٢١٢) * · الا أن القاعدة العامة المطردة فى الاشتقاق بالنسبة لمثل هذه الكلمات المشتقة هو العودة الى أصل الحروف فى الثلاثى · وعلى هذا ففى حالة كلمة « قيمة » نعود الى الأصل « قوم » مرة أخرى ناظرين الى الواو مرة أخرى فنقول « قوم » و « تقويما » ·

ولكن لوحظ أن بعض العرب أهملوا النظر الى أصل الحرف فى الجذر الثلاثى ونظروا الى حالته الراهنة بعد ابداله فى الكلمة ، وهذا ما يسميه بعض النحاة « الاشتقاق على التوهم » ، أو ما يسمه بعض المحدثين «الاشتقاق من المشتق ، دفعا للغموض وازالة للبس ·

ومن هذه القاعدة الخاصة أجاز مجمع اللغة العربية فى القاهرة أن

(*) يشير الرقم الاول الى رقم المرجع فى القائمة المثبتة فى نهاية كل باب من أبواب الكتاب ، بينما يشير الرقم الثانى بعد العلامة (:) الى رقم الصفحة أو الصفحات فى ذلك المرجع نفسه · وعندما توجد علامة الوصل (،) فان ما بعدها يشير الى رقم مرجع آخر وأرقام صفحاته وهذا هو النظام المتبع فى الكتاب كله ·

يقال قيمت الشيء تقييما بمعنى حددت قيمته . وذلك للتفرقة أو ازالة اللبس بين هذا المعنى وبين « قومته » بمعنى عدلته وجعلته قويما أو مستقيما ·

وهكذا توجد فى لغتنا المعاصرة كلمتان صحيحتان فصيحتان · ووجودهما على هذا النحو يمكن أن يحل لنا مشكلة التداخل فى المعنى والخلط فى الاستخدام حين نجدنا ازاء كلمتين أجنبيتين هما Valuation وفى رأينا أن أفضل ترجمة لها الآن هى « تقييم » ، والكلمة الأخرى هى Evaluation وأفضل ترجمة لها أيضا كلمتنا التقليدية المعتادة « تقويم » · فالاولى لا تتجاوز معنى تحديد القيمة أو القدر ، أما الثانية ففيها معنى التعديل والتحسين والتطوير ·

هل هو تقييم نفسى أو تقويم نفسى ؟

لا شك فى أن علم النفس يتضمن الكثير من عمليات « التقييم » بالمعنى الذى بيناه فى الفقرة السابقة ، أو بمعنى اصدار الأحكام على « قيمة » الأشياء أو الأشخاص أو الموضوعات أو الأفكار ·

فالانسان لا يكاد يتوقف عن التقييم ، واعطاء قيمة لما يدرك · الا ان هذا التقييم فى معظمه من النوع الذى يمكن أن نسميه « التقييم المتمركز حول الذات » ، ومعناه أن الشخص يحكم على الأشياء او الأشخاص او غير ذلك بقدر ما ترتبط بذاته هو · وقد يستخدم فى أحكامه هذه اعتبارات المنفعة أو الألفة أو الفائدة أو المكانة أو المركز الاجتماعى أو نقصان تهديد الذات أو سهولة الفهم والادراك · وهذه جميعا تصبغ الأحكام فى مثل هذه الحالة بصبغة ذاتية . وقد يكون بعضها أقرب الى التخمين لأننا قد نصيب فى أحكامنا هذه وقد نخطيء بنفس الدرجة من الاحتمال ·

وقد تكون أحكام الفرد ونتائج تقييمه للاشياء والأشخاص قرارات سريعة لا يسبقها فحص وتدقيق كافيين لمختلف جوانب الموضوع او الشخص الذى يصدر عليه الحكم ، وقد تتصف أحيانا بأنها لا شعورية ، بمعنى أن الشخص منا وهو يصدر هذه الأحكام لا يعى الدلالات والقرائن والأسس التى تعتمد

عليها أحكامه ، وقد لا يستطيع « التلفظ » بها مع أنه يكون على درجة كبيرة من الثقة فى صحتها ، وقد تكون بالفعل كذلك ، وهى بهذه الصفة تكون أقرب الى الأحكام الحدسية .

وقد يكون التقييم النفسى مرتكزا على أحكام تتوافر لدينا عنها دلالات واضحة فى الذهن ونكون على درجة كبيرة من الوعى بها ، معتمدين فى ذلك على فهم واف وكاف للظاهرة التى نقوم بتقييمها ، وعلى تحليل دقيق لعناصرها ومكوناتها . وتكون أحكامنا فى هذه الحالة ذات طبيعة استدلالية(*)

الا أن الأمر لا يتوقف فى بعض الأحيان عند هذا الحد . ففى المجال التربوى - مثلا - تتطلب دراسة السلوك الانسانى « عملية اصدار حكم على مدى تحقيق الأهداف التربوية ودراسة الآثار التى تحدثها بعض العوامل والظروف فى تيسير الوصول الى هذه الأهداف أو تعطيله» (١١: ٥٤٩) ومعنى ذلك أن اصدار الحكم فى مجال التربية يمكن أن يتبعه اجراء عملى يتعلق بتحسين العملية التربوية . وقد يكون ذلك فى صورة برنامج للتعلم العلاجى مثلا ، وهذا كله بالطبع يتجاوز معنى التقييم الى معنى التقويم .

وبالمثل ، حين يجرى الباحث فى المجال الكلينيكى دراسة على الطفل « للحكم » على مستواه العقلى ، من حيث التخلف أو التوسط أو التفوق . ان الأمريكون أدخل فى باب « التقييم النفسى » لو لم يتجاوز مرحلة اصدار الحكم على المستوى وهو من نوع تحديد القيمة أو القدر . وهذا أمر نادر فى الممارسة السيكولوجية المعتادة ، لأن الشائع أن يتبع ذلك « اجراء عملى » من نوع ما كالحاق الطفل بمدرسة للتربية الفكرية أو بفصل للمتفوقين ، أو تقديم نوع من التربية التعويضية . أو اعادة تنظيم البيئة التى يعيش فيها وغير ذلك. وكل هذا من نوع التقويم النفسى .

ومعنى ذلك أن التقويم النفسى هو الاجراء المعتاد فى المجالات النفسية والتربوية ، وهى المجالات التى شاع فيها مفهوم « التقويم » أكثر من غيرها .

(*) التمييز بين التخمين والحدس والاستدلال يستند الى تصور نظرى لاحد مؤلفى هذا الكتاب (راجع فؤاد أبو حطب المرجع رقم ١٠) .

صحيح أن المفهوم اتسع نطاقه فى السنوات الأخيرة وامتداد الى مجالات الاقتصاد والاجتماع والسياسة والطب وغيرها الى الحد الذى دفع بعض المهتمين الى الكتابة فيما يسمى « علم التقويم Evaluation Science » ، وفى جميع الحالات يتجاوز اللفظ حدود « تحديد القيمة » الى التحسين والتطوير والتنمية · ولعل هذا أقرب الى المعقول · فالباحث فى العلوم الانسانية - بصفة عامة - لا تتشابه مهمته مع مهمة الناقد الفنى أو الأدبى التى تقتصر على « وزن » العمل الفنى والأدبى « بميزان النقد » ليحدد قيمته وجدواه ، ولا تمتد بالطبع الى اقتراح الطرق التى يمكن بها أن يتحسن العمل الفنى أو الأدبى أو يتطور ، وهو عمل أقرب الى التقييم منه الى التقويم ·

والباحث فى علم النفس ليس « متفرجا » على السلوك الانسانى ولا تتحدد أهدافه فى مجرد الحكم عليه لتحديد قيمته أو قدره أو مقداره أى تقييمه وانما تمتد فى كثير من الأحوال الى تقويم هذا السلوك بالمعنى الذى أشرنا اليه ، أى أن التقويم النفسى هو القاعدة العامة والتقييم النفسى هو الاستثناء·

ومن قبيل هذا الاستثناء ما يحدث فى مجال التربية حين تقتصر مهمة المدرسة فيما يتصل بالامتحانات على الحكم على المتعلم بالنجاح أو الفشل ·

ومعنى ذلك أن مهمة الاخصائى النفسى اقرب الى مهمة الطبيب ، التى لا تقتصر على مجرد « قراءة » ميزان حرارة الجسم أو مقياس ضغط الدم أو « مشاهدة » الرسم الكهربائى للمخ أو القلب ، لمجرد الحكم بالسواء أو المرض . وانما تتجاوز ذلك الى التشخيص ثم العلاج ·

التقويم النفسى والقياس النفسى :

من المفاهيم التى يحدث بينها خلط كبير فى مجال علم النفس مفهوما التقويم النفسى والقياس النفسى ، وقد حددنا فيما سبق معنى التقويم النفسى، فماذا عن القياس النفسى ؟

يعرف نناللى (٢٧ : ٢٧) القياس فى العلم عامة وفى علم النفس خاصة بأنه « قواعد استخدام الأعداد بحيث تدل على الأشياء بطريقة تشير

الى كميات من صفة أو خاصية ٠ ومعنى ذلك أن القياس النفسى يعتمد فى جوهره على استخدام الأعداد ، الا أنه فى صورته المحكمة يتضمن فكـرة الكم والتى تعنى مقدار ما يوجد فى الشخص من خاصية معينة ٠

ورغم أهمية هذا التصور الكمى لخصائص الانسان الا انه ليس شائعا فى جميع الأحوال وبالنسبة الى جميع الخصائص ، فبعض الخصائص لايزال يستعصى على التناول الكمى المباشر وتستخدم فى وصفه « لغة كيفية » كاملة٠ وسواء لجأنا الى الوصف «الكمى» أو «الكيفى» للسلوك الانسانى فان ذلك يعد من قبيل البيانات التى لابد أن تتوافر للباحث النفسى قبل أن يصدر أحكامه أو يتخذ قراراته أو يقترح الاجراءات العلاجية والتصحيحية ٠ وعلى هـذا فان القياس النفسى يوفر للباحث البيانات والمعلومات الكمية التى يعتمد عليها عند القيام بعمليات التقويم النفسى، الا أنه ليس مصدر المعلومات الوحيد ، فاذا لم تتوافر لهذا الباحث الطرق الكمية لوصف السلوك ، أو اذا استعصت الظاهرة النفسية على التناول الكمى فلابد له من الاعتماد على الطرق الكيفية فى جمع المعـلومات ٠

ومعنى ذلك أن التقويم أكثر عمومية من القياس ، فالتقويم النفسى يعتمد على المقاييس أو على غيرها من طرق جمع المعلومات عن السلوك الانسانى٠ وفى كلتا الحالتين يتضمن التقويم اصدار الأحكام القيمية واتخاذ القـرارات العملية ٠ كما أن مقادير القياس الكمية أو الصـــيغ اللفظية الوصفية يمكن استخدامها استخدامات مختلفة تبعا لأهدافنا من عملية التقويم النفسى ٠ فنتائج الاختبار التحصيلى الواحد (والتى قد تتخذ صورة قياس كمى) يمكن استخدامها فى الحكم على مدى تحقيق هدف مثل الوصول بالتلاميذ الى المستوى المعتاد من الكفاءة ، كما يمكن استخدامها فى الحكم على مدى تحقيق هدف مختلف مثل تنمية قدرة معينة ، وبالطبع تختلف فى الحالتين خطـــة الباحث التى يستخدمها للوصول الى مثل هذه الأحكام وما يتـــلوها من خطوات عملية ٠

مستويات القيـــاس :

أشرنا الى أن القياس يستخدم لغة الكم ، الا أن هذا لا يعنى أن جميع

المقاييس من فئة واحدة، فقد قام علماء النفس بتحديد أنواع القياس المختلفة ودرجة ملاءمة العمليات الحسابية المعروفة لكل من هذه الأنواع ٠ ولهـــذا الموضوع أهميته القصوى لأننا لو لم نتناوله ببعض التعمق قد نفترض أن جميع العمليات الحسابية من جمع وطرح وضرب وقسمة يمكن أن تستخدم مع جميع نظم القياس ٠ بل قد يستنتج البعض أن القياس يصبح مستحيلا ما لم نستخدم جميع العمليات الحسابية ٠ ولو أخـــذنا ميدان القياس العقلى منذ بدايته لوجـدنا وعيا مبكرا بهـــذه المسألة فعلى ســـبيل المثـــال نجـــد أنـــه منـــذ ظهـــور الاختبارات المبكرة فى هـــذا المجال أشـــار العلماء أنه لا معنى بأن الطفل الذى نسبة ذكائه ٦٥ يكون نشاطه العقلى نصف طفـل آخر نسبة ذكائه ١٥٠ ، فالأطفال يختلفون دون شك كثيرا فى استجابتهم للمواقف التى تتطلب الاستدلال المجرد ، الا أنه لا يوجد فى سلوكهم ما يبرر التعبير عن هذه الفروق فى صورة معامل مقدار ٢ أو كسر مقداره ½ أو نسبة مقدارها ٢ : ١ ، بل لا معنى مطلقا لقسمة نسبة ذكاء على نسبة ذكاء أخرى ٠

وحينما تنبه علماء النفس الى هذه الخاصية التى تتوافر فى الأعداد التى يستخدمونها تحققوا من أنه توجد مقاييس مشابهة خارج علم النفس ٠ فالحرارة مثلا تقاس فى العادة بمقاييس مئوية (أو فار نهايتية) فاذا انخفضت الحرارة من ٣٠ درجة مئوية أثناء النهار الى ١٥ درجة مئوية أثناء الليل لا نستنتج من ذلك أن الجو أصبح فى منتصف الليل نصف دفئة أثناء الظهيرة. وذلك لأن الصفر فى مقاييس الحرارة صفر اعتباطى ولا يعنى « عدم وجود حرارة على الاطلاق » . كما أن درجات الحرارة فوق الصفر لا يمكن تناولها بنفس الطريقة التى نتناول بها درجات الطول أو الوزن ٠ ومن الواضح أن مقاييس الذكاء أقرب الى مقاييس الحرارة منها الى مقاييس الأطوال ٠

والواقع أن أفضل تصنيف لأنواع القياس المختلفة أو مستوياته وبخاصة فى ميدان علم النفس ذلك اقترحه ستيفنس (٣٠) وفيه يقسم الطرق المختلفة لاستخدام الأعداد الى أربعة أنواع هى المقاييس الاسمية ومقاييس الرتبـــة والمسافة والنسبة لكل منها قواعده وحدوده وضوابطه ، ولكل منها الاجراءات الاحصائية الملائمة له ٠ وقد طور كومبس (٢٣) هذا التصنيف وحدد العلاقات التى تربط بين المستويات المختلفة وأضاف فئات جديدة نعرضها فيما يلى :

١ ـ **المقايس الاسمية** : هذا النوع من المقاييس يقع فى أدنى مستويات القياس فى ضوء مدى ملاءمة الاجراءات الحسابية المعروفة له ومدى تطبيقها

فيه • وتسمى المقاييس الاسمية nominal كذلك لأن الأعداد تخصص لتعيين عناوين lables أوفئات categories ينتمى اليها الأشخاص أو الأشياء كأفراد •

ومن أمثلة استخدام الأعداد لتدل على عناوين ما يقوم به الاخصائى الجيولوجى حين يختار عددا من عينات الصخور ويعطيها ١ ، ٢ ، ٣ • الخ ، وفى هذه الأحوال تستخدم الأعداد كما تستخدم العناوين ، ولا يتضمن هذا بحال من الأحوال أى مطلب من التحليل الرياضى • ومن هذه الاستخدامات أيضا تلك الأرقام التى تخصص لكل لاعب من فريق كرة القدم ، او الأرقام على علامات الطريق أو أرقام العناصر الذرية •

ويجب أن تؤكد أن أى مقياس يهتم باستخدام مقصود للاعداد • ومن هذه الاستخدامات المقصودة التسمية أو العنونة labeling • وفى هذه الأحوال لا يوجد أى غرض لاستخدام التحليل الرياضى للاعداد ، كما ان الأعداد لا تعد دالة على كميات من خصائص •

ويرتبط باستخدام الأعداد كعناوين استخدامها لتدل على مجموعات أو فئات من الأشخاص أو الأشياء • ومن أمثلة ذلك اعطاء الرمز (١) ليدل على الاناث والرمز (٢) ليدل على الذكور عند التصنيف حسب الجنس ، ويمكن أن تستخدم الأعداد فى أى تصنيف آخر للاسوياء والمضطربين ، او حسب المهن ، أو حسب الحالة الاجتماعية ، الخ • والفرق الوحيد بين استخدام الأعداد فى التسمية واستخدامها فى التصنيف هى أننا فى حالة التصنيف يتم تجميع أكثر من وحدة بعضها مع بعض ولذلك فان الوحدات التى تعطى نفس العدد تتشابه فى خاصة من الخصائص • وكما هو الحال بالنسبة لاستخدام الأعداد فى التسمية فان الأعداد التى تستخدم لتدل على فئات فى نظام معين للتصنيف ليس لها أى مضمون كمى • وفى كلتا الحالتين فان العملية الحسابية التى يمكن تطبيقها على المقاييس الاسمية هى عملية العد أو التعداد للحالات الفردية ، أما الأعداد المستخدمة ذاتها فلا يمكن ان تستخدم معها عملية الجمع أو الطرح أو الضرب أو القسمة •

ويرى كومبس أن هذا النوع من المقاييس تحكمه علاقة التساوى ،

وهذا يعنى أن أى زوج من الأشياء يجب أن ينتمى بوضوع الى نفس الفئة او لا ينتمى اليها · وعادة ما يحكم علاقة التساوى هذه مبدأ الانتظام symmetry بمعنى أنه اذا كانت ا = ب ، .·. ب = ا ، ومبدأ التعـدى transitive بمعنى أنه اذا كانت ا = ب ، ب = حـ ، .·. ا = حـ · وبالطبع تلعب العوامل الثقافية دورا هاما فى تحديد مثل هذه الفئات الاسمية وابتكار فئات جديدة ·

٢ ـ مقاييس الترتيب الجزئى : يرى كومبس (٢٣ : ٤٧٤ ـ ٤٧٥) أنه فى الفئة الواحدة من فئات المقاييس الاسمية قد نجد ما هو اكثر من مجرد أن الوحدات التى تتألف منها الفئة متساوية فيما بينها ومختلفة عن الوحـدات التى تتألف منها فئة أخرى · فقد توجد بعض العلاقات بين بعض وحـدات الفئات المختلفة ، ومن هذه العلاقات أن وحدات احدى الفئات قد تكون اكبر (أو أصغر) من وحدات فئة أخرى · فاذا كان لدينا عدد من الفئات المتكافئة وكانت هذه العلاقات موجودة بين كل فئتين منهما فاننا نحصل على مقياس الترتيب الجزئى partial ordinal .

ولكى نوضح ذلك لنفرض أننا نريد قياس ما يسمى المستوى الاقتصادى ـ الاجتماعى ، ولنفرض أيضا على سبيل التبسيط ان هذه الخاصية تتألف من مكونين هما مستوى الدخل والمستوى التعليمى ، فاننا فى هذه الحـالة اذا وجدنا أن الفرد (ا) اعلى دخلا من الفرد (ب) ، وأنه فى نفس الوقت أرقى منه فى المستوى التعليمى يمكننا أن نقول أن (ا > ب) فى المسـتوى الاقتصادى والاجتماعى ، وكذلك اذا كان (ب) اعلى فى مكونى الخاصـية من فرد ثالث هو (حـ) فان (ب) لا يصبح فقط أعلى من (حـ) فى المستوى الاقتصادى ـ الاجتماعى ولكن (ا) يصبح أيضا أعلى من (حـ) فيه أى (ا > حـ) · ومن هذا يتضح أن العلاقة هنا متعدية مثل النوع الأول من المقاييس ولكنها ليست منتظمة asymmetric لأننا فى المثال السـابق نقول أنه اذا كان ا أكبر من ب فان ب ليس اكبر من ا ·

لنفرض أيضا أن لدينا شخصا رابعا هو (د) مستواد فى مكـونى الخاصية اقل من (ا) وأعلى من (جـ) · اننا فى هذه الحالة نستطيع القول أن ا > د > حـ · ولكن لنفرض فى نفس الوقت أن بالرغم من أن (د) أعلى

دخلا من (ب) الا أنه أقل منه تعليما · اننا فى هذه الحالة تواجه مشكلة حقيقية لأننا لا نستطيع أن نحدد مباشرة ما اذا كان (ب> د) أو (د > ب) بالنسبة للمستوى الاقتصادى والاجتماعى · وهكذا يصبح (ب) ، (حـ) لايمكن المقارنة بينهما ويتخذ مقياس المستوى الاقتصادى الاجتماعى للافراد الأربعة (ا ، ب ، حـ ، د) صورة الترتيب الجزئى على النحو التالى :

ا

ب   د

حـ

حيث يدل على أن الشخص الأعلى له مكانة أكبر من الشخص الأدنى، وحيث لا يوجد اتصال بين (ب ، د) اللذين لا يمكن المقارنة بينهما · ويمكن الحصول على هذا النوع من المقاييس حين تتألف الخاصية أو السمة من عنصرين أو أكثـــر لا يرتبطان فيما بينهما بعـــلاقة الاضافة additive أو التعويض Compensative .

٣ ـ **مقاييس الرتبة** : المستوى التالى من مستويات القياس هو ما يسميه ستيفنس مقايس الرتبة ordiral scales ، ويستخدم هذا النوع حين يمكن تنظيم الوحدات فى سلسلة تمتد بين الأدنى والأعلى فى الخاصية التى نقيسها، الا أننا لا نستطيع أن نحدد بدقة الفرق بين أى رتبتين · والواقع أن الترتيب هو أكثر صور القياس بدائية (باستثناء المقاييس الاسمية) · وهذا يعنى انـــه أساس لكل الصور الراقية من القياس ، ومع ذلك لا يحمل الا القليل من المعلومات عن الخاصية التى نقيسها · فحين يقوم الفاحصون فى موقف المقابلة بترتيب مجموعة من المتقدمين لشغل احدى الوظائف فانهم يستخدمون مقياس الرتبة ، كما أن المئينيات التى تستخدم كثيرا فى تفسير نتائج الاختبارات هى نوع من الترتيب وتحمل خصائص مقياس الرتبة · وفى هذا النوع من المقاييس لا يمكن استخدام العمليات الحسابية المعروفة من جمع وطرح وضرب وقسمة ، ولا يمكن استخراج متوسط الرتب أو انحرافها المعيارى ، وانمـــا تصلح لهذه المقاييس أساليب احصائية أكثر ملاءمة · وهى من النوع الذى يسمى الاحصاء اللابارامترى ومن أشهرها معامل ارتباط الرتب ·

ويمكن أن يتحول مقياس الترتيب الجزئى الى مقيـــاس رتبة كامل اذا أمكن مساواة عناصر الخاصية بعضها ببعض · ففي مثال قياس المســتوى الاقتصادى ـ الاجتماعى الذى أشرنا اليه يمكننا أن نساوى كل عام من التعلم بجنيه واحد من الدخل الشهرى (١٢ جنيها فى السنة) ، مثلا · وفى هذه الحالة نحصل على أساس اجرائى بسيط للمقـــارنة بين أى شـــخصين أو تصنيفهم فى فئات متساوية · ويصبح المقياس فى هذه الحالة مقياسا للرتبة مادامت عناصر السمة تخضع للتحويل الخطى البسيط بحيث يحل أحـــد العناصر محل الآخر على أساس قاعدة أو مبدأ مفهوم ·

٤ ـ **مقياس الرتبة المترى** : من الملاحظ على أنواع المقـــاييس الثلاثة السابقة أن عناصر المقياس هى فئات من الأشخاص أو الأشياء ، وأن العلاقات بينها هى علاقات التساوى أو « أكبر من » ، ولا يتضمن مفهوم المسافات بين الفئات · ومعنى ذلك أننا قد نلاحظ أن (أ) أكبر من (ب) وأن (ب) أكبر من (حـ) الا أننا لا نستطيع أن نحدد ما اذا كانت المسافة بين (أ) و (ب) أكبر أو أقل من المسافة بين (ب) و (حـ) · ولهذا يرى كومبس (٢٣) أن المقاييس السابقة جميعا هى من قبيل المقايس الاسمية لفقدانها خاصية هامة فى القياس عامة وهى خاصة تحديد المسافات بين فئات الأشياء أو الأشخاص·

الا أنه بالنسبة لجميع المقاييس السابقة لو أمكننا تحديد المسافات بين الفئات فاننا نحصل على مستوى أعلى من القياس ، فمثلا اذا كانت العلاقة (أكبر من) تعود على بعض المسافات بين الأشياء المتجاورة فى مقياس رتبة ، فان هذا المقياس يصبح مقياس رتبة مرتبا جزئيا ، أما اذا كانت هذه العلاقة تصدق على جميع المسافات يصبح المقياس مرتبا ترتيبا كاملا ويطلق كومبس(٢٣) على المقياس فى هاتين الحالتين اسم مقياس الرتبة المترى Ordered metric scale ، ويقصد به المقياس الذى يمكن أن نجد فيه ثلاثية من الفئات أ > ب > جـ > د.

والمثال التالى يوضح ذلك : لنفرض أننا نريد قياس سمة « السلطة » كما تتمثل فى سلوك « الرئاسة » فاننا نستطيع ذلك فى ضوء العلاقة « أكبر من » التى أشرنا اليها فى حديثنا عن مقياس الرتبة المعتاد ، فاذا كان (أ)

يرأس (ب) ، (حـ) يرأس الجماعة التى ينتمى اليها (أ) فاننا نقول ان حـ > أ · وهكذا عن طريق تتبع التسلسل الرئاسى يمكن اعداد مقياس رتبة ·

الا أن مقياس الرتبة البسيط هذا يمكن تحويله الى مقياس رتبة مترى اذا توافرت لدينا معلومات تحدد لنا الفروق فى مقدار السلطة الذى يمارسه الشخص مقارنا بآخر · ولكى نوضح ذلك نعطى مثالا من الميدان التربوى · لنفرض أن الموجه الفنى يرأس اثنين من المدرسين الأوائل ، وأن كلا من هذين المدرسين يرأس ١٠ مدرسين عاديين ، فاننا نستنتج فى هذه الحالة أن سلطة الموجه تزيد على سلطة المدرس الأول بمقدار ٢ مدرس أول ، ٢٠ مدرسا عاديا · فاذا انتقلنا فى سلم السلطة من الموجه العادى الى الموجه الأول الذى يرأس ٣ موجهين عاديين ، يرأس كل منهم ٢ من المدرسين الأوائل اللذين يرأس كل منهما أيضا ١٠ مدرسين عاديين ، فان سلطة الموجه الأول فى هذه الحالة تزيد على سلطة الموجه العادى بمقدار ٤٠ مدرسا عاديا (٣٠ + ١٠)، ٤ مدرسين أوائل ، ٢ موجهين عاديين · وهكذا بالنسبة لكل مستوى من مستويات الرئاسة التربوية · ومن هذا المثال يتضح أن الموجه الأول يفوق الموجه العادى فى السلطة بمقدار أكبر مما يفوق به الموجه العادى المدرس الأول ، ونستنتج أن الزيادة فى مقدار السلطة من المدرس الأول الى الموجه العادى أقل من الزيادة من الموجه العادى الى الموجه الأول ·

٥ ـ **مقاييس المسافة** : يمكن القول أن ما يميز مقياس المسافة interval عن مقياس الرتبة بمستوياته المختلفة التى اشرنا اليها أنه يسمح بتحديد مدى بعد شيئين أو شخصين بعضهما عن بعض فى الخاصية موضوع القياس ، وأن تكون هذه المسافات متساوية ويحتاج ذلك الى وضع قواعد معينة يتم الاتفاق عليها لاستخدام الأعداد حتى يمكن تحديد كم الصفة أو الخاصية فى الشىء أو الشخص · ومن أمثلة ذلك أننا يمكننا أن نحصل على مقياس مسافة للاطوال فى جماعة من الأطفال اذا لجأنا بدلا من قياس الطول مباشرة ـ الى قياس طول كل طفل بالنسبة لأقصر طفل فى المجموعة · وبالتالى فان أقصر طفل فى هذه المجموعة يحصل على الدرجة صفر ، والطفل الأطول منه بخمسة سنتيمترات يحصل على الدرجة ١ ، والأطول بعشرة سنتيمترات يحصل على الدرجة ٢ ، وهكذا · والاجراء الأقرب الى الشيوع فى أغلب المقاييس النفسية أن تحدد المسافات فى ضوء بعد كل طفل عن المتوسط

الحسابى لأطوال المجموعة بمقادير معيارية كما تتمثل فى الانحراف المعيارى· لنفرض أن المتوسط ١٢٠ سم والانحراف المعيارى ٥ سم ، فاننا نستطيع فى هذه الحالة أن نعطى + ١ للطفل الذى يزيد طوله عن هذا المتوسط بمقدار ٥ سم ، والدرجة ــ ١ للطفل الذى يقل طوله عن المتوسط بمقدار ٥ سم أيضا · وهذه المسافات حول متوسط الأطوال وغيره من الخصائص يمكن حسابها دون معرفة حقيقية بمدى بعد الأشخاص عن نقطة الصفر ·

والواقع أن أغلب المقاييس النفسية والتحصيلية والعقلية من هـذا القبيل ، فنحن نقارن درجات طالبين فى الاختبار ونحدد مدى بعد كل منهما عن المتوسط · الا أن أهم نواحى القصور فى هذه المقاييس انها ليس لهـا صفر مطلق · فقد يحصل التلميذ على الدرجة صفر فى احـد الاختبارات التحصيلية الا أن ذلك لا يعنى أنه لا يوجد لديه معلومات على الاطلاق حول موضوع الاختبار ، كما أن الطفل الذى أعطيناه صفرا فى الطول فى مثالنا السابق لا يعنى أنه ليس له طول على الاطلاق وانما الصفر فى هذه الأحوال هو صفر اعتباطى ·

ويمكن أن تستخدم مع مقاييس المسافة عمليات الجمع والطرح ، الا أن عملية القسمة بالذات لا يجوز استخدامها على الاطلاق · فلا نستطيع أن نقسم الدرجة التى حصل عليها التلميذ (أ) فى الاختبار على الدرجة التى حصل عليها التلميذ (ب) فى نفس الاختبار ، لأن القسمة تفترض مقدما وجود الصفر المطلق ، وبالتالى لا يمكن الحصول على النسـبة (كالضعف أو النصف) · ولتوضيح ذلك نضرب المثال التالى :

نفرض اننا طبقنا اختبارا على شخصين فحصل الأول على الدرجة ٨٠ وحصل الثانى على الدرجة ٤٠ ، ولنفرض أن الباحث الذى أعد هذا الاختبار ضمنه بالصدفة ١٠ مفردات أخرى يسهل على كل من هذين الشخصين الاجابة عليها اجابة صحيحة ، ففى هذه الحالة تصبح درجة الشـخص الأول ٩٠ درجة والشخص الثانى ٥٠ وفى هذه الحالة يظل الفرق بين الدرجتين ثابتا ، أى ٤٠ ، فى الحالتين ، ولكن معامل الدرجتين لن يكون متساويا ، فبدلا من أن يكون ٢ (أى ٨٠ ÷ ٤٠) فى الحالة الأولى يصبح ١ر٨ (أى ٩٠ ÷ ٥٠) فى الحالة الثانية · وهكذا فاننا فى الاختبار الواحد لا توجد لدينا طريقة

لايجاد ما اذا كانت معلومات احد الأشخاص ضعف معلومات شخص آخر أو نصفها أو ثلاثة امثالها · ولكننا حين نفترض أن كل مفردة مؤشر للمعلومات يتساوى فى جودته مع أى مفردة أخرى ، فاننا لا نخرق مبادىء الرياضيات أو المنطق حين نطرح درجة الشخص الأول من درجة الشخص الثانى أو حين نجمع هذه الدرجات ونستخرج متوسطها وانحرافها المعيارى · والواقع أن الطرق المختلفة التى تستخدم فى ترجمة الدرجات الخام فى الاختبارات الى الأنواع المختلفة من المعايير ـ وهى الطرق التى سنشير اليها فيما بعد ـ تعتمد فى جوهرها على عمليتى الجمع والطرح ·

٦ ـ مقاييس النسبة : تعد مقاييس النسبة ratio scales أعلى مستويات القياس وفيها تقاس خاصية معينة فى الشىء بوحدة أو مقدار معيارى من هذه الخاصية بهدف معرفة عدد الوحدات المعيارية من هذه الخاصية التى توجد فيه · فمثلا عندما نقيس طول الحجرة بهدف عد الأمتار أو السنتيمترات (وحدات القياس) التى توجد وتتكرر فى هذا الطول ، فاننا نستخدم هذا النوع من المقاييس ·وتصلح مع مقاييس النسبة كل العمليات الحسابية من جمع وطرح وضرب وقسمة ومع صلاحية هذه العمليات يمكن استخدام الرياضيات العليا كالجبر والهندسة التحليلية والتفاضل والتكامل وكل الطرق الاحصائية المتقنة · وتتوافر فى مقاييس النسبة جميع خصائص مقاييس المسافة بالاضافة الى الصفر المطلق · ونحن نألف هذا النوع من المقاييس أكثر من غيره لأن جميع الأبعاد الفيزيائية المعروفة كالطول والوزن والحجم يمكن قياسها بهذه الطريقة · ولذلك نستطيع القول أن الشخص الذى طوله ١٨٠سم ضعف الشخص الذى طوله ٩٠ سم · والواقع أن تسمية هذا النوع من المقاييس باسم مقاييس النسبة جاءت من قابلية هذه المقاييس للقسمة أو التعبير عنها فى صورة نسبة ·

وهذا النوع من المقاييس لا يوجد الا قليلا فى علم النفس ، ولا تتوافر المقاييس القليلة من هذا النوع الا حين نقيس الخصائص العقلية أو الوجدانية بواحدات فيزيائية ، كأن نقيس زمن الرجع أو التعلم بوحدات زمنية (كالثانية أو أجزاء الثانية) ·

طرق القياس فى علم النفس :

يهتم القياس فى العلم دائما بجانب معين من جوانب الأشياء ، أو خاصية من خصائصها • وعلى وجه التحديد نستطيع القول اننا لا نقيس الأشياء وانما نقيس خصائصها ، ويصدق هذا على الانسان ، فنحن لا نقيس الطفل مثلا وانما نقيس ذكاءه أو اجتماعيته • ولهذا التمييز أهميته من ناحيتين : أحدهما أن القياس يتطلب عملية تجريد • فالخاصية تدل على علاقة بن الأشياء فى بعد معين : كالوزن أو الذكاء ولا تختلط بالجوانب (أو الأبعاد) الأخرى للاشياء وهذه المسألة رغم وضوحها للقارىء المتخصص الا أنها لا تبدو كذلك لغير المتخصصين ، فكثيرا ما نجد خلطا بين خاصية معينة للاشياء وكل الخصائص التى يمكن ملاحظتها فيها ، والفشل فى تجريد خاصية معينة يجعل مفاهيم القياس صعبة الفهم ، ومن ذلك مثلا يصعب على كثير من الناس أن يفهم أن الجانح والسوى يمكن أن يتوافر لهما نفس المستوى من الذكاء (كما يقاس باختبارات الذكاء) •

والسبب الثانى للتأكيد على أن القياس يتناول خاصية معينة هو أن ذلك يجبرنا على ضرورة دراسة طبيعة الخاصية قبل محاولة قياسها لان من المحتمل ألا توجد هذه الخاصية • فليس من الضرورى أن جميع العبارات التى نستخدمها فى وصف الناس تقابلها خصائص قابلة للقياس •

ويوجد احتمال آخر أن القياس قد يتناول خليطا من الخصائص لا خاصية واحدة • وهذا ما نجده كثيرا فى استخبارات « التوافق » أو قوائم مشكلات الشباب والتى تتضمن مفردات ترتبط بعدد من الخصائص المنفصلة • ورغم أن هذه المقاييس المختلطة لها ما يبررها من الوجهة العملية والتطبيقية الا أنها لا تتضمن الا قليلا بمنطق القياس فى العلم •

والخصائص التى تخضع للقياس فى العلم قد لا تقاس مباشرة بوحدات معيارية منها ، فالحرارة تقاس بمقياس مترى ، وكذلك الوزن ، كما أن معظم هذه الخصائص من نوع التكوينات الفرضية hypothetical constructs التى لا تلاحظ هى فى ذاتها مباشرة ، أى أنها لا تخضع للادراك الحسى المباشر • وكثير من المفاهيم الكبرى فى العلم من هذا القبيل ، فالمغناطيسية

والنشاط الذرى وانتقال الحرارة والذكاء من نوع التكوينات الفرضية أو التجريدات المعرفية . وبالطبع تحتاج دراستنا لهذه الظواهر وقياسها توافر قدر من الوقائع التى يمكن ملاحظتها أو المؤشرات الخاصة بها . واذا كان علماء الفيزياء يعتمدون فى دراستهم لظواهر المغناطيسية والكهرباء والحرارة على مؤشرات مثل تغيير اتجاه ابرة البوصلة أو نشاط عداد جايجر أو قراءة الترمومتر ، فان علماء النفس فى دراستهم وقياسهم لتكويناتهم الفرضية كالذكاء ، والقدرة الميكانيكية والميل العلمى والاتجاه نحو العمل اليدوى وغيرها يعتمدون على مؤشرات على هذه الخصائص النفسية تنتمى الى عالم الواقع ويمكن ملاحظتها ملاحظة موضوعية مباشرة هى أساليب الأداء performance والتى تتخذ احدى صور ثلاث :

(أ) الأداء اللغوى كما يتمثل فى النطق والتحدث والتلفظ شفاها وقد يمتد الى وسائل الاتصال غير اللفظى كالايماءات والاشارات وغيرها .

(ب) الأداء الحركى كما يتمثل فى نشاط أعضاء الحركة مباشرة كاستخدام الجسم أو الأيدى أو الأصابع أو الأقدام .

(حـ) الأداء الفسيولوجى كما يتمثل فى نشاط الأجهزة الجسمية المختلفة كالنشاط الهورمونى أو نشاط القلب أو نشاط المخ .

ويرتبط مفهوم القياس بالعد . فان ما يفعله المرء فى أى مقياس هو عد الوحدات المتشابهة . فيقاس الطول بعدد الوحدات المتساوية من الأمتار (السنتيمترات مثلا) ، ويقاس الوزن بعدد الوحدات المتساوية فى صورة جرامات (مقدار جذب الأرض لسنتيمتر مكعب من الماء النقى) . ولا شك أن الخصائص الفيزيائية لبعض أدوات القياس لها فائدتها فى أنها تقوم بعملية العد ذاتها بدلا من أن يقوم بها الملاحظ بنفسه . وترتبط المقاييس النفسية أيضا ارتباطا وثيقا بالعد . وبالتالى فان الكرونسكوب يفيد كعداد لأجزاء الثانية فى قياس زمن الرجع . وبعد عد أزواج الكلمات التى يتذكرها المفحوص فى تجربة عن تعلم الارتباطات الثنائية يصل الفاحص الى تحديد مقدار ما تم تعلمه . وبالمثل فاننا بعد عد الأسئلة التى أجاب عليها المفحوص فى اختبار معين نصل الى قياس لقدرته .

ويعد أداء الفرد بعد عرض قائمة الارتباطات الثنائية ، أو حله للسؤال فى الاختبار مؤشرا على الخاصية النفسية موضع القياس •

ويستخدم علماء النفس عددا من الطرق فى قياس الخصائص النفسية توضح لنا كلها مقدار السمة أو الخاصية في ضوء مؤشراتها ، وأهم هذه الطرق ما يأتى :

١ ـ تكرار أو احتمال حدوث مؤشر السمة أو الخاصية :يمكن القول أن عدد الاستجابات « الصحيحة » فى اختبار للقدرة مثال لما نسميه تكرار حدوث مؤشر السمة ، لأن كل مفردة من مفردات الاختبار هى « فرصـــة » للدلالة على ما اذا كانت استجابة المفحوص تدفعه الى الطـرف الاعلى من المقياس أو تهبط به الى طرفه الأدنى أو تضعه فى المستوى المتوسط •

فمثلا اذا أجاب المفحوص اجابة صحيحة على ٦٠ سؤالا من ١٠٠ سؤال فان قياس السمة بدلالة تكرار حدوث مؤشرها يمكن أن يتحول فى هذه الحالة الى احتمال فى صورة نسبة مئوية فنقول أن احتمال حدوث مؤشر السمة فى هذه الحالة هو ٦٠٪ •

فاذا كان العمل الذى يؤديه المفحوص يتالف من وحـــدات (أسئلة) كثيرة متساوية فى الصعوبة أو السهولة فيمكن أن تقاس السمة بعـــدد الاستجابات الصحيحة (سواء كانت فى صورة تكرار أو احتمـــال) التى يؤديها المفحوص فى وقت محدد ، وفى هذا نحن نقيس السرعة ، وقد تقاس السمة بعد الاستجابات الخاطئة التى تصدر عن المفحوص فى وقت محدد وفى هذا يصبح الباحث مهتما بقياس الدقة •

٢ ـ شدة حدوث مؤشر السمة أو الخاصية : وتتمثل طريقة الشدة أو الحـــدة intensity فى اختبارات القدرة فى مستويات صعوبة المفردات التى يستطيع المفحوص الاجابة عليها ، وتتمثل فى النشاط العضلى بمقدار الطاقة أو الجهد المبذول كما يقاس بسعة الاستجابة magnitude أو ارتفاع وعلو الاستجابة altitude أو قوة الاستجابة strength ، وفـــى بعض صور النشاط الأخرى فان المكونات الانفعالية مثل ضغط الدم أو معدل

النبض أو توصيل الجلد للتيار الكهربائى أو التوتر العضلى تدل على الشدة · وتدل شدة الاستجابة بالقبـــــول أو الرفض للقضايا الخـــلافية على قوة الاتجاهات ·

وتستخدم هذه الطريقة فى القياس حين يكون على الباحث المبالغة فى تأكيد القوة · ويمكن أن تستخدم كمؤشر للسمة حين يحصل الفاحص على مقياس يتألف من أسئلة مرتبة حسب الصعوبة · ويصبح السؤال فى هذه الحالة ما هو الحد الذى يصل اليه المفحوص ولا يتعداه ؟ ويمكن اعطـــاء مثال واضح من ميدان الألعاب الرياضية وخاصية فى قفز الحواجز حيث يرفع الحاجز تدريجيا حتى يصل اللاعب الى الحد الذى لا يستطيع اجتيازه ، ويمكن الحصول بهذه الطريقة على مقياس للأداء · ويصدق هذا على مقياس بينيه للذكاء الذى يتألف من أسئلة متدرجة فى الصعوبة ، ويتطلب هذا جهدا كبيرا فى تحديد هذه المستويات · وقد تعطى للأسئلة درجـــات متناسبة مع صعوبتها بالنسبة للمفحوصين وتصبح درجة المفحوص هى مجموع هـــذه الدرجات الموزونة للأسئلة التى يجب عليها اجابة صحيحة أو خاطئة (حسب نظام وزن الدرجات) وقد يعطى للمفحوص درجة رتمة لأصعب أو أســـهل سؤال أجاب عليه ·

٣ ـ مدى حدوث مؤشر السمة أو الخاصية : وهذه الطريقة ليســـت واضحة أو شائعة الاستخدام كالطريقتين السابقتين وتتطلب جوهرها تحديد عينة متنوعة من الأسئلة يتألف منها الاختبار ويستخدم فى قياس السمة تنوع الأسئلة التى أجاب عليها المفحوص كما يتحدد مثلا بعدد فئات الاستجابة ·

تعريف الاختبار :

تعرف أنستازى (١٨) الاختبار النفسى أنه « مقياس موضـــوعى مقنن لعينة من السلوك » · ويعرفه كروتباك (٢٤) بأنه طريقة منظمة لمقارنة سلوك شخصين أو أكثر · وقد ناقش أحد المؤلفين هذين التعريفين فى موضـــع آخر (١٠) وتوصل منهما الى تعريف أكثر شمولا · ونبدأ بعبـــارة « عينة من السلوك » التى وردت فى تعريف أنستازى فنقول أن الاختبارات النفســـية تتشابه مع الاختبارات التى تستخدم فى العلوم الأخـــرى ، بشرط أن تتم

الملاحظات على عينة صغيرة جيدة الانتقاء من سلوك الفرد . وفى هذا الصدد فان المتخصص فى علم النفس يحذو حذو المتخصص فى علم الكيمياء مثلا حين يختبر دم المريض أو ماء الشرب فيقوم بتحليل عينة او اكثر منه . فاذا أراد السيكولوجى أن يختبر المحصول اللفظى للطفل . أو قـــدرة موظف الحسابات على أداء العمليات الحسابية أو مدى التآزر بين العين واليد عند الطيار ، فانه يفحص أداء هؤلاء فى « عينة » ممثلة من الكلمات او العمليات الحسابية أو الأعمال الحركية . أما مسألة شمول الاختبار السلوك موضوع الاختبار أو عدم شموله فتتوقف على وحدات العينة وطبيعتها . فالاختبار الذى يتكون من خمس مسائل فقط ، أو من أسئلة فى الضرب فقط يعد مقياسا غير جيد للمهارة العددية . واختبار المفردات اللغوية الذى لا يتضـــمن الا مصطلحات الألعاب الرياضية لا يعطينا تقديرا يعول عليه للمحصول اللغوى للطفل .

وتعتمد القيمة التشخيصية diagnostic والتنبؤية predicive للاختبار على مدى تقنينه وموضوعيته الا أن الموضوعية والتقنين هما من شروط الاختبارات الجيدة الى جانب شروط أخرى سنتناولها بالتفصيل فيما بعد . وتضمينها فى التعريف يخل بشرط الايجاز اللازم . الا أن هـذه القيمـة التشخيصية والتنبؤية للاختبار تعتمد فى جوهرها على مدى استخدام الاختبار كمؤشر على مجال أوسع وأكثر أهمية من السلوك . فمن النادر أن تكون عينة السلوك التى يشملها الاختبار هدف القياس النفسى فى ذاتها وانما المهم أن تبرهن على وجود قدر من التطابق الوثيق بين عينة الســلوك فى الاختبار والمجال الأكثر شمولا الذى ينتمى اليه (بين معرفة الطفل لقائمة الكلمات فى اختبار الاستعداد اللغوى واتقانه الكلى للمفردات اللغوية مثلا كما تتمثل فى قراءاته لاحدى مجلات الأطفال) .

ويجب أن نلاحظ فى هذا الصدد أن الأمر لا يحتاج الى وجود التشابه التام بين أسئلة الاختبار والسلوك الذى تسعى الى تشخيصه أو التنبؤ به . وانمـــا الضــــرورى هـــو البرهنـــة على وجـــود تطـــابق تجـــريبى بينهمــا (وهـــذا ما يســـمى الصـــدق) . ولذاك تتفــاوت درجـــة التشــابه بين عينـــة الاختبـــار وميدان السلوك الذى تنتمى اليـــه . ففى الأحوال المتطرفة قد يتطابق الاختبار تطابقا تاما مع السلوك الذى يتنبأ به أو يشخصه ، ومن ذلك اختبارات التحصيل المدرسى أو الكفاءة المهنية . وتوجد

درجة أقل من التشابه تتمثل فى كثير من اختبارات الاستعدادات المهنية ٠ وفى الطرف الآخر نجد الاختبارات التى لا ترتبط ارتباطا مباشرا بميـــدان السلوك ومنها الاختبارات الاسقاطية ٠ ورغم هذه الاختلافات السطحية فان جميع هذه الاختبارات تتألف من عينات من السلوك ، ومن الواجب اثبات جدواها بالبرهان التجريبى على وجود تطابق بين أداء المفحوص فى الاختبار وغير ذلك من المواقف ٠

الا أن الاختبار قد لا يكون بالضرورة عينة سلوك فى كل الحالات ٠ ففى ميدان الاختبارات التحصيلية يزداد الاهتمام فى السنوات الأخيرة بما يسمى الاختبارات المرجعة الى المحك criterion-referenced tests وتقوم فى جوهرها على مسلمة ملاحظة معظم أو جميع السلوك وليس عينة منه وخاصة حين يكون موضوع الملاحظه وحدة متكاملة من وحدات التعلم المدرسى ، وفى هذه الحالة يصنع الاختبار التحصيلى بحيث يشمل ما تم تدريسه فى هذه الوحدة من محتوى وأهداف وليس عينة منهما ٠ وفى ضوء نتائج هذه الاختبارات المرجعة الى المحك يمكن تشخيص صعوبات التعلم والتأكد من مدى الاتقان . كما يمكن اقتراح الطرق العلاجية لتحقيق أهداف التعلم التى يمكن أن يحرزها اكثر من ٩٠٪ من التلاميذ ٠ وفى هذه الحالة تصبح العملية تقويما وليس تقييما كما هو الحال فى معظم الممارسات التى تعتمد على ما يسمى الاختبارات المرجعة الى المعيار ٠

واستخدام كرونباك لعبارة « طريقة غير منظمة » فى تعريفه للاختبار اكثر دقة من كلمة « مقياس » عند انستازى ، لأن فى قول انستازى نوع من عدم التمييز بين المصطلح « اختبار » test والمصـــــطلح « مقياس » measure ، فعلى الرغم من تداخل معانيهما الا أنهما ليسا مترادفين ٠

فمن ناحية نجد أن لفظ « مقياس » أكثر عمومية لأنه يستخدم فى كل ميادين البحث السيكلوجى عندما نسعى الى الحصول على أوصاف « كمية » كما هو الحال فى بحوث الادراك والاحساس والحكم والمجال السيكوفيزيائى العام ٠ أى أن اللفظ يستخدم فى الأغراض السيكولوجية العــامة ، بل وفى صميم علم النفس التجريبى ٠ فكثيرا ما نقيس التعلم أو الاستجابة أو المثير ، وتستخدم فى هذه الأغراض المقاييس الفيزيائية أو النفسية ٠

ويطلق على المقياس لفظ اختبار فى مجال استخدامه فى ميدان علم النفس الفارق وحده · وعلى هذا فان مقاييس العتبات الفارقة أو التعلم أو الادراك يمكن أن تستخدم « كاختبارات » اذا تحول اهتمامنا بها الى ميدان الفروق الفردية · الا أن الاختبار يتكون فى العادة من عدد من الأســئلة أو المفردات التى لا تأخذ صورة مقاييس النسبة وانمــا قــد تكون من نوع مقاييس المسافة أو الرتبة ، ومعنى ذلك انه ليست جميع المقاييس اختبارات الا عند الأهتمام بعلم النفس الفارق · وفى هذه الحالة فقط يمكن أن يحـل لفظا اختبار ومقياس كل منهما محل الآخر ·

ومن ناحية أخرى ليست جميع الاختبارات مقاييس · وقد اشرنا الى معانى القياس ومستوياته وهو فى كل الأحوال يتطلب نوعا من الوصف الكمى سواء كان من نوع الكم المتصل أو الكم المنفصل · وليست جميع الاختبارات من هذا القبيل · فقد نجد بعض الاختبارات التى لا تعطى درجة للمفحوص وانما يستخدمها الأخصائى النفسى لمساعداته على الوصول الى وصف لفظى أو كيفى للمفحوص (مثل طرق الملاحظة) · ولا يتطلب الأمر فى هذه الأحوال استخدام المقاييس من أى مستوى من المستويات التى ذكرت آنفا ·

ومادامت الاختبارات هى فى جوهرها أدوات الدراسة العلمية للفروق الفردية فانها تسعى الى المقارنة بين الأفراد كما يقول كرونبــاك فى تعريفه الا أن هذه المقارنة لا تكون بين الأفــراد فى ضــوء معيار norm فحسب ، وانما قد تكون أيضا داخـــل الأفـــراد فى ضـــوء مســـتوى standard أو محـــك criterion ، كما أن هذه المقارنة لا تكون فى عينة السلوك فقط ، كما هو الحال فى الاختبارات المرجعة الى المعيار ، وانما تشمل أيضا المقارنة فى « معظم » السلوك أو « كله » كما هــو الحال فى الاختبارات المرجعة الى المحك ·

والمعيار هو أساس للحكم على أداء المفحوصين والمقارنة بينهم فى ضوء أدائهم الفعلى ، ويأخذ الصيغة الكمية فى أغلب الأحوال ، ويتحدد فى ضوء الخصائص الواقعية لهذا الأداء ، ومن ذلك استخدام المتوسط الحســابى لدرجات عينة التقنين معيارا لوصف الأداء العادى فى الاختبار ، وفى ضوئه

تتحدد الأوضاع النسبية للافراد فنقول أعلى من المتوسط او اقل من المتوسط او متوسط ·

اما المستوى فيتشابه مع المعيار فى انه اساس للحكم على الأداء فى ضوء هذا الأداء ذاته ، الا أنه يختلف عنه فى جانبين : أولهما انه قد ياخذ الصورة الكمية أو الكيفية ، وثانيهما أنه يتحدد فى ضوء ما يجب ان يكون عليه الأداء وليس ما هو عليه بالفعل · ومن هذه المستويات ما نجده فى نظم الامتحانات المعتادة حين نقارن درجات التلاميذ فى هذه الامتحانات بنظام النهايات الصغرى والكبرى ، أو حين تتحدد تقديرات النجاح قبليا فى صورة ضعيف ومقبول وجيد وممتاز فى ضوء نسب مئوية من النهاية العظمى للمادة الدراسية توضع مقدما ولا تحسب بالطرق الاحصائية فى ضوء الأداء الفعلى فى الامتحانات · وهذه جميعا وسائل غير دقيقة فى تحديد المستوى ، أما أفضل الطرق فتكون حين يقارن الأداء كما يحدده الاختبار بمستوى الجودة أو الاتقان أو التمكن الذى يحدده الهدف التربوى أو التعليمى أو المهنى ، ويكون تحديد هذا المستوى فى الأصل قد تم فى ضوء ما يجب أن يكون عليه الأداء ·

أما المحك فهو أساس خارجى مستقل للحكم على الأداء فى الاختبار ، وقد تكون هذه المحكات كمية أو كيفية ، فمثلا لكى نحكم على نجاح برنامج تعليمى أو تدريبى فى تحقيق أهدافه يمكن مقارنة أداء التلاميذ فى الاختبارات التحصيلية المرتبطة بهذا البرنامج بمستويات الكفاية الانتاجية التى تتحدد فى الميدان الفعلى للعمل ·

وهكذا توصل فؤاد أبو حطب (١٠) الى تعريف يقترحه لمفهوم الاختبار وفيه يقول :

« الاختبار النفسى هو طريقة منظمة للمقارنة بين الافراد أو داخل الفرد الواحد فى السلوك أو فى عينة منه فى ضوء معيار أو مستوى أو محك » ·

انواع الاختبارات :

يمكن أن نقسم الاختبارات على أسس كثيرة منها الشكل أو الغرض أو المحتوى ويقسمها كرونباك قسمين :

١ ـ اختبارات الأداء الأقصى maximum performance التى تستخدم اذا كنا نريد أن نعرف الى أى حد يستطيع الشخص ان يقوم بأداء ما الى أقصى قدرته ، ويمكن الاشارة اليها باسم اختبارات القدرة ·

٢ ـ أما القسم الثانى فيتضمن الاختبارات التى تهدف الى تحديد الأداء المميز typical performance وهى نقيس ما يحتمل أن يفعله الشخص فى موقف معين أو فى نوع معين من المواقف ، وتدخل فى هذا النطاق اختبارات الشخصية والميول ·

١ ـ اختبارات الأداء الأقصى (أو اختبارات القدرة) :

أهم ما يميز هذه الاختبارات أن الشخص يشجع فيها ليحصل على احسن درجة يمكنه الحصول عليها ، وتشمل اختبارات الذكاء واختبارات القدرات المختلفة سواء كانت عقلية أو حركية · ومن هذه الاختبارات مجموعة كبيرة تسمى مقاييس القدرة العقلية العامة وهى تهدف الى قياس تلك القدرات العقلية العامة فى معظم انماط التفكير والتعلم (وهى بهذا المعنى اختبارات ذكاء) · وهناك اختبارات القدرات الأكثر تخصصا وهى القدرات ذات القيمة فى مهام محدودة ، منها مثلا الفهم الميكانيكى ، والاحساس بالدرجة الصوتية والمهارة الأصبعية · وهذه تسمى اختبارات القدرات المتخصصة · وهذه الاختبارات قد تستخدم منفردة أحيانا كما قد تستخدم مجتمعة ، وقد تستخدم مجموعة منها تقيس قدرات خاصة متعددة لتساعد فى رسم بروفيل لقدرات الفرد ·

أما اختبارات الكفاءة proficiency فتقيس القدرة على أداء مهمة او عمل ما له أهميته فى ذاته : مثل قراءة الفرنسية · أو العزف على البيانو او اصلاح خلل فى ماكينة طائرة · وحيث أن هذه الاختبارات تقيس أداء

أشخاص تدربوا على العمل المقاس فانها تسمى احيانا اختبارات التحصيل achievement .

وهناك اختبارات الاستعدادات aptitude وتستعمل هذه الاختبارات للتنبؤ بالنجاح فى مهنة أو فى تدريب معين · فهناك مقاييس للاستعداد الهندسى مثلا ، أو الاستعداد الموسيقى أو الاستعداد للجبر وهكذا· ولا تختلف هذه الاختبارات من حيث الشكل عن الأنواع الأخرى · فاختبار الاستعداد الهندسى مثلا قد يتضمن اقساما تقيس القدرة العقلية العامة · وأخرى تقيس الفهم الميكانيكى أو المكان واقساما تقيس التحصيل فى الرياضيات · ويسمى الاختبار على أى حال ، اختبارا تحصيليا اذا كان يقيس نجاح الشخص فى الدراسة السابقة ·ويسمى اختبار استعداد عندما يستعمل للتنبؤ بنجاحه فى عمل المستقبل ·

٢ ـ اختبارات الأداء المميز :

لا تستعمل اختبارات الأداء المميز لمعرفة ماذا يستطيع الشخص ان يفعله بل لمعرفة ما يعمله بالفعل وطريقة عمله (مثلا فى الاختبار لادارة عمل ليس المهم ما يمكن أن يعمله الشخص المرشح ولكن ما يعمله بالفعل وأسلوبه فيه)· وهناك بعض الخصائص الشخصية · مثل البشر او الأمانة أو سعة الأفق لا تستطيع اختبارات القدرة أن تعرفنا بها ، على الرغم من أهميتها ، فالأداء المميز أو السلوك الفعلى هام من غير شك اذا نحن أردنا أن نقدر الاستعداد للنجاح فى عمل ما ·

واذا كانت الدرجة العالية شيئا مرغوبا فيه فى اختبارات القدرة ، ففى اختبارات الأداء المميز لا تستطيع أن تقول ان استجابة ما فى اختبار من اختباراتها « جيدة » ، اذ ليس هناك ما هو « جيد » أو « ردىء » مثلا فى أن يكون عند شخص ما ميل للهندسة · وكذلك نجد أن بعض الناس يغلب عليهم سلوك السيطرة (وهذه صفة من صفات الشخصية) ، ونحن لا نستطيع أن نحدد أن درجة معينة من السيطرة هى الأنسب والأفضل لأن العالم يتسع لأناس مختلفى الأنماط والصفات ·

ونستطيع القول بصفة عامة أن سلوك الشخص المميز له هو أفضل مفتاح لشخصيته ، فالعادات السلوكية لها قيمة تنبؤية فى ذاتها ، فما فعله الشخص من قبل يكون أميل الى أن يعيده ويكرره فى مواقف تالية فيما بعد ، ولكن يعترض معظم علماء النفس على افتراض أن عادات الشخص المنظورة هى شخصيته ، ذلك أن مواقف جديدة تنشأ دائما ، وان أى وصف للعادات المألوفة عند شخص ما لا تدل بصورة مباشرة على ما سيفعله فى موقف جديد . فربما عرف عن شاب أنه يكره البنات ثم تأتى بنت بعينها فتأسره بحبها ويدله بها . والمعالج النفسى الذى يستطيع أن ينشىء علاقة متينة وعميقة وينجح مع كثير من عملائه قد لا يستطيع أن ينشىء مثل هذه العلاقة مع شخص معين لأنه يثير عنده - أى عند المعالج - الكره والغضب مثلا . ولكن مع هذا ، ينبغى أن لا تنسينا هذه الاستثناءات أو الظروف الشاذة أن السلوك وردود الأفعال للمواقف المختلفة تعتبر حقا انعكاسات لتركيب شخصية أساسى ومتسق ، وهذا التركيب الأساسى والمتسق ينبغى استنتاجه من السلوك ، ويأمل العالم النفسى أنه عندما يفهم تركيب الشخصية هذا سوف يستطيع التنبؤ بأستجابات الشخص وسلوكه حتى فى المواقف الجديدة .

ويعتبر قياس الأداء المميز أمرا شاقا الى درجة بعيدة ، وقد حقق النجاح بنسب متفاوته . وبطرق متعددة . وهذه الطرق يمكن تقسيمها الى قسمين . الأول : ملاحظات السلوك ، والثانى : وسائل التقرير الذاتى . أما ملاحظات السلوك فهى محاولات لدراسة الشخص وهو فى حالة سلوكه الطبيعى . والملاحظات يمكن أن تتم فى موقف اختبار مقنن أو فى موقف غير مقنن أو طبيعى . أما الملاحظة المقننة فتتطلب أن يوضع كل شخص فى نفس الموقف الذى يوضع فيه الآخرون ، فيمكن أن تلاحظ « الشخصية » فى موقف اختبار عقلى ، أو أثناء مناقشة جماعية . كما انه من الممكن أن تبتكر مهام أو أعمال تتيح مجالا طيبا للملاحظة . وتسمى هذه المهام عندئذ اختبارات الأداء الشخصية .

والملاحظة المقننة تسمح بمقارنة شبه تامة لأشخاص لا يمكن أن يلاحظوا عادة فى ظروف متشابهة . هذا فضلا عن أنها تكشف عن خصائص لا نشاهدها الا عرضا فى حياتنا اليومية . وقد استعملت هذه الطريقة مثلا فى ملاحظة ردود الفعل المميزة للحبوط وذلك بأن يتاح للشخص أن يبدأ عملا

ولكن يحال بينه وبين أن يحقق غرضه أو يصل الى هدفه بشكل ما ويعطى رد فعله أو الطريقة التى يسلك بها فى موقف الحبوط ذاك صورة واستبصارا فى مدى ضبطه الانفعالى •

واذا كان للملاحظة أن تظهر السلوك المميز فعلا فلابد ان لا يعلم الشخص الخاصية أو الصفة التى يراد ملاحظتها أو تجرى ملاحظتها فعلا ، ففى موقف الحبوط المشار اليه منذ قليل من الممكن أن نذكر للشخص أو الأشخاص الذين نختبر ضبطهم الانفعالى ، اننا نقيس ذكاءهم • على اننا ينبغى أن نؤكد عند هذه النقطة أن هذا لا يجوز الا فى مجال البحث العلمى والتجريب ، ولابد أن يذكر للاشخاص الذين تجرى ليهم التجربة الهدف وما تم فيها بعد أن تنتهى التجربة وكذلك يجوز استخدام هذه الطريقة فى مجال الاختبار لعمل ما لمعرفة مدى توفر صفة أو صفات معينة مناسبة لعمل معين •

وكذلك يمكننا أن نحصل على معلومات عن السلوك المميز بملاحظة عينات من سلوك الشخص العادى فى نشاطاته اليومية فى الميدان كما هى ، كأن نلاحظ الأطفال فى الملعب حيث تظهر عاداتهم وشخصياتهم ، وكذلك الموظفين فى مكاتبهم ، أو العمال فى مصانعهم ، وهكذا • وطريقة الملاحظة الميدانية هذه يمكن أن تستخدم فيها أدوات تسجيل مقننة (واعلى هذه الأدوات من حيث تقنينها التصوير السينمائى) ، أو يستخدم فيها التسجيل والحكم الانطباعى •

أما القسم الثانى من طرق قياس الأداء المميز فهو وسائل التقرير الذاتى والأساس الذى تقوم عليه هذه الوسائل هو أنه لما كان الشخص تتاح له فرصة كبيرة ليلاحظ سلوكه الذاتى ، فانه يستطيع اذا أراد ، أن يعطى تقريرا له قيمته عن سلوكه المميز ويساعد فعلا فى التعريف بهذه الناحية • وتستخدم الأستخبارات أو الاستبيانات questionnaires كوسيلة من وسائل التقرير الذاتى عن الأداء المميز • على أن المشكلة الرئيسية فى التقرير الذاتى ، اذا كنا سوف نفسره على أنه صورة للسلوك المميز ، هى الأمانة ، فاذا حاول الشخص أن يعطى أفضل صورة ممكنة عن نفسه بدلا من وصفها كما هى فان الاختبار يفشل فى تحقيق غرضه ، وحتى عندما يحاول أن يكون صادقا فلن يخلو تقريره عن ذاته من خطأ تحريف أو تحيز •

ومعظم استخبارات التقرير الذاتى تعطى صورة تكاد تكون كاملة عن الشخصية ، وبعضها متخصص فى موضوعه ، فهناك مثلا استخبارات قياس عادات الدراسة ، واستخبارات الميول ، واستخبارات الاتجاهات الاجتماعية، وهكذا · وهناك نوع آخر من استخبارات التقرير الذاتى تسمى استخبارات التكيف أو اختبارات الخلق · وهذه الأسماء تدل على الطريقة التى تفسر بها درجات الاختبار ولكنها لا تميز شكلا معينا للاختبار ·

ومن المتفق عليه بصفة عامة ألا تستعمل استخبارات الشخصية كلمة اختبار فى عناوينها ، وذلك منعا لسوء استعمال غير المدربين على اعطاء هذه الاختبارات وتفسيرها وسوء فهم لوظيفة الاستخبارات فى التقويم النفسى ·

البرنامج التقويمى :

البرنامج التقويمى هو خطة منظمة للحصول على المعلومات والبيانات الصحيحة عن الأفراد موضوع التقويم ، ولكى نوضح طبيعة البرنامج التقويمى نعطى مثالا من المجال التربوى ، حيث يعنى البرنامج التقويمى مجموعة ادوات القياس والتقويم التربوية والنفسية الشاملة والمتنوعة التى تطبقها مجموعة من المختبرين والمرشدين النفسيين والتربويين والمشرفين الاجتماعيين والمدرسين ، وفق نظام مرسوم ، يحقق حصول المؤسسة التربوية على معلومات وبيانات ترتب وتنظم بحيث يستطيع المشرفون على العملية التربوية أن يستعملوها فى اتخاذ انسب القرارات للمتعلمين وللمؤسسة التربوية نفسها · اى أن البرنامج التقويمى يشمل : أدوات قياس ، ومختبرين ، وخطة للاختبار والتقويم ، وبيانات منظمة ، وطرق الاستفادة من هذه البيانات ، ثم علاقات اجتماعية وتفاعل اجتماعى بين هيئة التقويم وبين سائر اعضاء المؤسسة التربوية ، وبينها وبين مؤسسات اجتماعية اخرى فى المجتمع تقتضى طبيعتها وطبيعة العملية التقويمية الاتصال بها ·

ويجب أن تكون نظرتنا الى البرنامج التقويمى نظرة صحيحة فيكون العمل فيه محققا لأهدافه واغراضه ، والنظرة الصحيحة الى البرنامج التقويمى هى التى تعتبره متكاملا مع النظام التربوى كله ، كما تعتبره جزءا من التعليم ومن التخطيط التربوى · أى أنه لا يصح أن ننظر الى البرنامج

التقويمى على أنه شيء منفصل عن تيار العمل التربوى فى مؤسساته المختلفة· أو منفصل عن عملية التعلم، أو خارجها أو فوقها· ولن يتحقق نجاح لبرنامج تقويمى تربوى لا يقوم على هذا الأساس الكبير : **التكامل مع النظام التربوى ومع عملية التعليم ذاتها** ·

والمسئول عن وضع برنامج تقويمى وتسييره ، فى أى مستوى من مستويات العملية التربوية ، فى مدرسة أو معهد أو كلية أو منطقة تعليمية أو قطر بأسره ، يمكن أن يسير وفق خطوات عامة فى بنائه للبرنامج التقويمى على أن تهديه فى خطواته نظرة صحيحة الى التقويم ·

أما هذه الخطوات فيمكن تلخيصها فيما يلى :

١ ـ **تحديد الأهداف** : على هذا المسئول أن يبدأ بتحديد أهداف العملية التربوية العامة التى يضعها المجتمع ويتفق عليها والتى تحددها السلطات التربوية المسئولة التى تعبر عن أهداف هذا المجتمع وحاجاته ، ثم عليه أن يحدد الأهداف التربوية الخاصة بالمؤسسة التى سيتولى اقامة برنامج تقويمى فيها · وبعد أن يحدد هذه الأهداف العامة والخاصة عليه أن يقوم بتعريف القائمين على العمل التربوى فى المؤسسة بهذه الأهداف ، ويشاركهم الرأى والمناقشة فى مبرراتها وتفاصيلها ووسائل تحقيقها · ولا بأس بأن يوضع دليل مطبوع عن البرنامج التقويمى وعلاقته بهذه الأهداف يفيد منه المدرسون والآباء ومن يعنيهم الأمر ·

٢ ـ **جمع البيانات** : وبعد تحديد الأهداف والتعريف بها ، على المسئول عن التقويم أن يبدأ عملية اختيار الاختبارات التى يمكن الحصول عليها والتى تناسب كل هدف من هذه الأهداف ، مراعيا تنوع هذه الاختبارات وتمثيلها لكل الجوانب التى يراد قياسها · ومن الضرورى مراعاة الشروط الأساسية فى اختيار الاختبارات والتى سنتناولها بالتفصيل فيما بعد ·

أما اذا لم تكن هناك اختبارات أو مقاييس مناسبة لأهداف معينة ، فعلى القائم على البرنامج التقويمى بالاشتراك مع هيئة مساعدة له من المتخصصين فى القياس ، ان وجدت ، أو من المهتمين فى المؤسسة التربوية،

ان يبدأ فى وضع اختبارات ووسائل قياس خاصة ومناسبة · ويمكن بطبيعة الحال الاستعانة ببعض العلماء المتخصصين فى القياس النفسى فى وضع هذه المقاييس ، وأن يراعى فى جميع الأحوال شروط بناء هذه الأدوات ·

والهدف من اختيار الاختبارات أو بنائها هو استخدامها فى جميع البيانات اللازمة للتقويم ، وقد يحتاج الأمر الى بيانات اخرى لا توفرها الاختبارات مثل الوثائق والسجلات المدرسية ونتائج الامتحانات السابقة ، وهذه البيانات وغيرها لابد من توافرها حتى يكون الحكم على مدى تحقيق الأهداف التربوية سليما ·

وعند أعطاء الاختبارات وتطبيقها لابد من انتقاء العينات موضع البحث انتقاء سليما ، كما لابد أن تراعى الشروط الخاصة باعطاء الاختبارات وبهذا تبدأ عملية الحصول على المعلومات والبيانات عن الجوانب التربوية المختلفة موضع التقويم ·

وبعد الحصول لى هذه البيانات من مصادرها الصحيحة لابد من تنظيمها بطريقة تجعل الاستعانة بها والاستفادة منها مسالة يسيرة تساعد فى اتخاذ القرارات الملائمة له ·

وقد يحتاج الأمر الى تنظيم برنامج للتدريب الدورى لمن يقومون باعطاء الاختبارات على الأساليب الجديدة وعلى الاختبارات الجديدة التى يمكن الاستفادة منها فى البرنامج التقويمى ، ويدخل فى هذا تدريب المدرسين الذين يبدون اهتماما خاصا على استعمال وتطبيق بعض الاختبارات والطرق المستحدثة ·

٣ - **اصدار الأحكام** : فى ضوء المقارنة بين البيانات التى يحصل عليها المسئول من التقويم باستخدام الأدوات الملائمة والأهداف التربوية المحددة يتم اصدار الأحكام ، وبالطبع فان هذه الأحكام قد لا تكون مهمة خبير التقويم منفردا ، وانما قد يشاركه فيها عدد من المسئولين فى المنظومة التعليمية ، ومنهم المتعلمون أنفسهم · وتزداد هذه المقارنة يسرا اذا كان الهدف قد صيغ فى عبارات اجرائية ، لأن هذه الصيغة تتضمن تحديدا لمستوى

الجودة أو الاتقان أو التمكن ، بالاضافة الى أنها تسهل لنا اختيار الأدوات التى ترتبط ارتباطا مباشرا بالهدف ٠

٤ - **اجراءات التطوير** : تعد خطة التطوير والتنمية والتحسين والعلاج فى ضوء الأحكام التى يتم الوصول اليها ، وقد تشمل قرارات التطوير تقديم برنامج فى التربية التعويضية أو التعلم العلاجي أو ايداعا فى احدى المؤسسات أو التدخل بأساليب الارشاد النفسى أو العلاج النفسى المختلفة ٠

خصائص البرنامج التقويمى :

هناك سمات وخصائص لابد أن تتوفر فى أى برنامج تقويمى تربوى حتى يحقق الأهداف المرجوة منه يؤدى وظيفته بنجاح (٦ : ٣٨ - ٤٦) ٠

وأهم هذه الخصائص :

١ - **الشمول** : يجب أن يراعى البرنامج التقويمى الشمول ، وبيان هذا أنه لا ينبغى أن ينحصر الاهتمام فى القياس على المعارف والحقائق والمفاهيم٠ بل يتسع ليشمل الاتجاهات والميول والتفكير الناقد والتوافق الشخصى والاجتماعى ، وأن تستعمل شتى الأساليب بتوازن ، والحق انه اذا كان اهتمام المشرف على عملية التعليم يهتم بالتلميذ اهتماما كليا شاملا ، فمن اللازم أن يكون اهتمام البرنامج التقويمى أيضا شاملا لكل جوانب التلميذ وانشطته ٠

٢ - **التنظيم** : يجب أن تجمع نتائج الاختبارات وبياناتها وتنظيم وتحلل بحيث يصبح تفسيرها ممكنا وذا مغزى ، فنتائج البرنامج التقويمى سواء كانت كمية ، أى فى صورة درجات أو مذكرات صغيرة أو ملاحظات عابرة ، أو كيفية ، يجب أن تلخص فى اطار واضح من التقديرات وأن تتحول الى صورة احصائية أو رسوم بيانية أو تقارير لغوية ٠ وأن توصف اتجاهات هذه النتائج باستخدام الطرق الملائمة من التحليل الاحصائى أو تحليل المحتوى ، وأن تقدم فى جميع الأحوال فى صورة قابلة للفهم والنقل والاتصال بحيث تعطى صورة عن الفرد يمكن مقارنتها بصورة سابقة له ، فيساعد هذا

على التعرف على اتجاهات نموه فى المجالات المختلفة ، وكذا مقارنتها بصور زملائه الآخرين ٠

٣ ـ **التكامل** : لابد أن تتكامل المعلومات فى استعمالها عند القيام بحكم معين ، وعندما نضع خطة عمل بناء على هذا الحكم ، فالمعلومات الخاصة بالنواحى الصحية والتوافق العاطفى والاجتماعى والاتجاهات وغيرها يجب أن تربط مثلا بنتائج اختبارات التحصيل أو اختبار تحصيل بعينه ٠ وكل ما لدينا من بيانات ومعلومات عن الفرد لابد أن نعامله باعتباره وحده يوضح ويكمل بعضها بعضا ٠

٤ ـ **الاستمرار** : يتميز البرنامج التقويمى الناجح بالاستمرار ، أى أنه فى المؤسسة التربوية الحديثة يعتبر عملية مستمرة ٠ فالملاحظات اليومية والتقديرات والاختبارات المستمرة هى التى تشكل العمليات التقويمية ، وعن طريقها يحاول المدرس أن يقوم نمو التلاميذ ويوجهه ٠ فلا ينبغى ان تكون هناك فترة محددة للتقويم ثم يتوقف بعدها أو ينتهى تقويم التحصيل المدرسى مثلا فى يوم حساب معلوم أو فى أيام معدودة يوضح فيها الميزان ٠ والحكم الذى ينطق به لا يفيد المتعلم بأن يدله على مواطن الضعف او القوة ٠ أو الصحة أو العلة فى أسلوب التعلم والتحصيل والمراجعة ، فلا يفيدمن هذا الحكم متعلم أو معلم وانما هو قول فصل ٠ أما النهج القويم فى تقويم التحصيل خاصة ، فهو أن يتداخل مع عملية التعلم نفسها ، يسبقها ، ثم اثناءها خطوة خطوة حيث يجدى القياس والمقارنة ويكون حقا قولنا أن التقويم السليم ضرورى للتعلم السليم ، بل وضرورى لفاعلية المؤسسات التربوية على اختلاف أنواعها ومستوياتها ، ويتمثل ذلك خاصة فى الاهتمامات الحديثة بما يسمى التقويم التكوينى ٠

٥ ـ **الوظيفية** : لابد أن يعرف كل أعضاء المؤسسة التربوية التى يعمل فيها البرنامج التقويمى بما يمكن أن يقدمه لهم هذا البرنامج من خدمات ٠ فيعمل مثلا على أن يعرف التلاميذ والمعلمون ، بل وأولياء الامور ، أنه من الممكن اعطاء اختبارات خاصة غير تلك التى تطبق على كل التلاميذ ، او يمكن اعطاء اختبارات اضافية اذا أراد التلميذ ان يتعرف على بعض جوانب شخصيته ٠ وربما كان أخذ مثل هذه الاختبارات بابا يدخل منه التلميذ الى

مناقشة بعض مشكلاته فيما يتصل بعلاقته بزملائه أو بمدرسيه أو بعمـــله المدرسى أو حياته الأسرية ·

البرنامج التقويمى اذن دعامة رئيسية من دعائم العملية التربوية ، ويعتمد نجاحها الى حد كبير على نجـــاحه ، لهذا كان من الواجب على المسئولين عن العملية التربوية أن يعطوا البرامج التقويمية التربوية كل جهد أو مال تستحقه ، وهو جهد ومال مبذول فى خير وجه ، وانه لمعط ثمـــاره أضعافا مضاعفة ·

دور الاختبارات فى البرنامج التقويمى :

على الرغم من أن دور الاختبـــارات فى البرنامج التقويمى يبدو لنا متواضعا ، حيث يحتل خطوة واحدة من خطوات هذا البرنامج ، الا انه فى واقع الأمر اكبر اثرا واشد خطرا من ذلك ، ونكاد نقول أن الاختبـــارات تكاد تكون القاسم المشترك بين معظم مكونات البرنامج التقويمى ابتداء من الأهداف وانتهاء بالأساليب التطويرية والعلاجية ، وفيما يلى نعرض لبعض الأدوار التى تقوم بها الاختبارات فى مختلف جوانب هـــذا البرنامج ويشمل ذلك : الأهداف واتخاذ القرارات والاجراءات التطويرية وخاصة الارشاد النفسى ·

أولا : دور الاختبارات فى صياغة الأهداف :

تساعد الاختبارات فى اختيار الأهداف التربوية وتوضيحها على نحـو غير مباشر ، فهى تحفز المؤسسة التربوية على صياغة اهدافها والتعبير عنها بوضوح فى عبارات سلوكية أو اجرائية ، الا ان هـــذا التحديد السلوكى الواضح فى ذاته ليس ضمانا لاختيار أهداف تربوية صالحة · فعملية اختيار الأهداف ليست من العمليات البسيطة التى تعتمد على مجرد التعرف أو الحكم المباشر، وانما هى من الأمور التىتتطلب المفاضلة بين انظمة قيمية ومعيارية· ومع ذلك فهناك مجموعة من المعلومات التى يمكن للمدرسة أو المعلم أن يحصل عليها مما يهيىء فرصا افضل لاختيار أكثر حكمة للاهداف · ومجمـوعة المعلومات هذه كما يصنفها رالف تايلر (٣٢) هى :

١ ـ مجموعة من المعلومات تتعلق بالأفراد انفسهم قدراتهم ومعلوماتهم (التقويم النفسى)

ومهاراتهم وميولهم واتجاهاتهم وحاجاتهم ، وبالطبع ترتبط هذه الخصائص بالمرحلة النمائية التى يمر بها هؤلاء الأفراد ومطالبها فى نواحى النمو الجسمى والفسيولوجى والعقلى والانفعالى والاجتماعى والمهنى · ومراعاة هذه الخصائص فى صياغتنا للاهداف يجعل عملنا التربوى معينا للفرد على تشخيص صعوباته وعلاج مشكلاته واشباع حاجاته ومواجهة مطالبه ·

٢ ـ مجموعة من المعلومات تتعلق بمطالب المجتمع · ويتطلب ذلك معرفة جوانب القوة والضعف التى يؤثر بها المجتمع الحديث على عمليات التربية والتدريب عموما بحيث يكتسب الفرد من العملية التربوية المعلومات والمهارات والقدرات والاتجاهات والميول التى لها أهمية فى المجتمع والتى تساعده على التكيف مع الثقافة وتحسينها ·

٣ ـ مقترحات المتخصصين فى مختلف ميادين المعرفة والتى يرون بها كيف يمكن لميادينهم أن تساهم فى عمليات تعليم واعداد وتدريب الأفراد ·

هذه الأنواع الثلاثة من المعلومات قد توحى بكثير من الأهداف التى قد تتعارض فيما بينها وبالتالى يصبح من الضرورى القيام بالاختيار من بينها وهذا الاختيار يعتمد أساسا على فلسفة المدرسة او المؤسسة التربوية او برنامج التدريب الذى نقوم به ، والتى يمكن بها تحديد الأهداف ذات القيمة الكبرى فى ضوء مفاهيم « المواطن » و « المجتمع » · كما يعتمد هذا الاختيار على نتائج الدراسات والبحوث فى سيكولوجية التعلم · وفى هذه الأحوال جميعا تلعب الاختبارات دورا هاما فى تزويدنا بالمعلومات حول الأفراد مما يوحى بأهداف معينة · وبعبارة أخرى يمكن القول بأن الاختبارات تلعب دورا مزدوجا فى هذه الخطوة الأولى من عملية بناء خطط التعليم والتدريب : تستثير البيئات التعليمية لاختيار وتحديد أهدافهم بوضوح حتى يمكن استخدام الاختبارات فى تقويمها ، كما أنها ذات فائدة كبيرة فى الحصول على المعلومات التى نستخدمها عندما نتخذ قرارا باختيار أهداف معينة دون غيرها ·

ويتوافر للاخصائيين فى التربية فى الوقت الحاضر مصادر هامة للاهداف قام باعدادها المهتمون بالتقويم والقياس النفسى والتربوى ، سواء

فى المجال المعرفى (٢٠) أو الوجدانى (٢٦) أو الحركى ، ويمكن الرجوع الى مصدر آخر للحصول على مزيد من التفاصيل (١١) ·

ثانيا ـ دور الاختبارات فى اتخاذ القرارات :

كل من يتعامل مع الناس يقوم باستمرار باتخاذ قرارات فيما يختص بسلوك هؤلاء الناس · وعملية اتخاذ القرارات بشأن الآخرين موجودة فى تعاملنا الاجتماعى بعضنا مع بعض ، حتى فى مواقف الحوار بين شخصين، حيث يتخذ أحد المتحدثين قرارا بتغيير الموضوع لأن الشخص الآخر بدا عليه من العلامات ما يوحى اليه بأن موضوع الحديث مؤلم له أو غير مرغوب فيه على أقل تقدير ، وكذلك بالنسبة للمواقف الاجتماعية الأخرى الأكثر تعقدا· كما أن عملية اتخاذ القرارات تحتل مكانا هاما فى الشئون الادارية وسياسة الأفراد حيث تقتضى طبيعة العمل اتخاذ قرارات فيمن يقبلون وفى توزيعهم على تخصصات مختلفة وترقيتهم بل والاستغناء عنهم · وكذلك نجد لاتخاذ القرارات أهميتها التى نحس بها جميعا فى مجال العملية التربوية ، فالمدرس يتخذ قرارا خاصا بمدى استعداد تلاميذه للانتقال الى موضوع دراسى جديد، والمدرسة تتخذ قرارات فيما يختص بتوزيع التلاميذ ، ونقلهم بل وفى معاملة من يصدر عنهم مشكلات سلوكية ملحوظة ، أو من يسعون الى حل مشكلاتهم لدى المتخصصين فى المدرسة مثل المشرف الاجتماعى ، أو الموجه التربوى ، أو رائد الفصل أو ناظر المدرسة أو أى مدرس يثق فيه التلميذ صاحب المشكلة·

وكلما كان لدى الشخص الذى يتخذ القرارات معلومات كافية وصحيحة الى حد بعيد عن الفرد أو الأفراد الذين يتخذ قرارا بشأنهم كانت الفرصة أمامه أرحب لأن يحصل على أطيب النتائج وأنسبها ، فتوفر المعلومات قبل اتخاذ القرار يساعد متخذ القرار على أن يصدر أحكم قرار ممكن ·

والاختبارات ، على اختلاف أنواعها نفسية أو تربوية أو طبية هى وسائل لاعطاء معلومات لا نستطيع أن نحصل عليها بغيرها، وهذه المعلومات تساعدنا فى اتخاذ القرارات ·

وأى قرار يتضمن تنبؤا ، ذلك لأن الاختبارات كلها تدلنا على بعض

الاختلافات فى أداء أفراد معينين لحظـة اجراءات الاختبارات عليهم ، ولكن هذه الحقيقة تكون عديمة القيمة اذا لم تستطع ان تتنبأ ان هؤلاء الأفراد يختلفون فى أداء آخر ، أو فى نفس الأداء ، فى وقت آخر • وللاختبـارات التى تتنبأ أهمية فى النواحى التعليمية ، وفى العلاج النفسى وفى البحث العلمى •

وتستعمل الاختبارات لاتخاذ قرارات من أنواع كثيرة ، منها : انتقاء الأفراد وتصنيفهم ، وتقويم طرق المعاملة المختلفة (طرق تدريس مثلا) ، وقبول أو رفض الفروض العلمية • وسوف نتكلم عن كل من هذه الأنواع من القرارات فيما يلى :

١ ـ **الانتقاء والتصنيف** : الانتقاء عبارة عن قبول أو رفض ، فعنـدما نعطى اختبارات متنوعة لمجموعة من الشبان المتطوعين للخدمة العسكرية أو المجندين ، ثم نقرر قبول بعضهم ورفض البعض الآخر حسب نتائج هذه الاختبارات المتنوعة ومستعينين بكل ما يمكننا الحصول عليه من معـلومات خاصة بهؤلاء الشبان ، يكون عملنا هذا انتقاء أما التصنيف فيقوم على تحديد المعاملة التى يتلقاها المتلقى ، فاذا نظرنا فيما سيتم بشأن الشبان المقبولين للعمل فى الخدمة العسكرية فنجد أن منهم من سيوزع على فرع الطيران ، أو الميكانيكا أو المشاة أو غيرها ويتلقى تدريبا أو معاملة مناسبة لهذا التصنيف، وكذلك نجد أن الشخص المصاب باضطراب انفعالى (أو مرض نفسى) يتحدد نوع العلاج الذى يتلقاه بناء على التشخيص المبنى على الاختبارات المختلفة، وهذا التشخيص عبارة عن تصنيف نوع الاضطراب أو المرض ، ويساعد التصنيف فى التركيز على نوع المعاملة العلاجية المناسبة لهذا الشخص والتى تؤدى الى أحسن النتائج معه •

ومن التصنيف أيضا تعيين الأفراد على « مستويات » للعمل وليس فقد لأنواع العمل • فاذا كان من نتائج اختبارات أعطيت للتلاميذ فى اللغة أن قمنا بتوزيعهم على مستويات تتراوح بين الضعيف والمتوسط والحسن ، ليتلقو تدريبا لغويا مناسبا ، كل حسب مستواه ، كان هـــذا تعيينا للافراد على مستويات للعمل •

ومن الضرورى أن نميز بين اختبارات الانتقاء واختبارات التصنيف . فاذا اعطينا اختبارات كانت نتيجتها استبعاد من عندهم اضطراب انفعالى مثلا كانت هذه اختبارات انتقاء ، لأن القرارات التى تبنى على المعلومات التى نحصل عليها منها قرارات قبول أو رفض فحسب ، لا تدلنا على أكثر من مدى الصلاحية أو عدم الصلاحية للقبول فى فئة أو مؤسسة ما ٠ أما الاختبارات التى تعطى معلومات تساعد على تحديد أو التعريف بطريقة معاملتهم أو تدريبهم أو علاجهم فهى اختبارات تصنيف ٠

ومن المهم أن نشير الى ضرورة الوصف الفردى المفصل فى عملية التصنيف ونعنى بهذا ضرورة الحصول على معلومات مفصلة قائمة على اختبارات متعددة ومن مصادر مختلفة عن الفرد ، من ناحية خصائصه الشخصية المميزة ، واتجاهاته وميوله ٠ واجتماعيته أو طبيعة علاقاته الاجتماعية ، الى جانب نتائج الاختبارات المحددة التى تهدف الى تصنيفه معين ، فمن اللازم أن تكون تحت أيدينا مثل هذه المعلومات عن الطالب الذى يرغب الدخول فى كلية الهندسة أو الطب الى جانب نتائج اختبارات الاستعداد ٠

(٢) **تقويم المعاملات** : اذا كانت للاختبارات أهميتها فى اتخاذ قرارات كما بينا منذ قليل ، فان لها أهميتها أيضا فى تقويم المعاملات ، أيا كان نوع هذه المعاملات تدريبا أو تدريسا أو علاجا ٠ ولعله يكون واضحا أن ما نعنيه بالمعاملة هنا هو الطريقة التى يعامل بها فرد أو أفراد أيا كان مجالها أو نوعها ٠ فالاختبارات التحصيلية مثلا ، أو اختبارات التحصيل المدرسى ، يمكن استخدامها لمقارنة طريقتين من طرق تدريس مادة الرياضة فى المدرسة الاعدادية للتعرف على أفضل الطرق فى تدريس هذه المادة ٠ ويمكن الحكم على سياسات الاشراف والأفراد ومقارنة فاعليتها عن طريق قياس الاتجاهات والروح المعنوية بين الأفراد ٠

(٣) **التحقق من الفروض العلمية** : كثيرا ما تستعمل الاختبارات النفسية بأنواعها المختلفة فى التحقق من الفروض العلمية التى نضعها فى مجالات علم النفس المختلفة ٠ والفرض العلمى بصفة عامة عبارة عن توقع معين أو تحديد لعلاقة وظيفية بين متغيرات ، ونستخدم الاختبارات للتحقق عن

مدى صحة هذا الفرض أو خطئه · فاذا اقمنا فرضا علميا يذهب الى أن الأطفال الذين لا ينشأون فى وسط عائلى طبيعى بين والديهم واخوتهم يتميزون بدرجة ما من الاضطراب الانفعالى بعكس الأطفال الذين ينشأون فى وسط عائلى طبيعى ، فان من الممكن التحقق من صحة هذا الفرض باعطاء مجموعتين من الأطفال ، مجموعة نشأت فى وسط عائلى طبيعى ومجموعة أخرى لم تنشأ فى هذا الوسط ، اختبارات شخصية تدلنا على درجة اتزانهم الانفعالى ونقارن نتائج المجموعتين بعضها بالبعض الآخر · وكذلك اذا كان لدينا فرض علمى يقول أن الأطفال يتعلمون بدرجة أحسن اذا تعاونوا فى مواقف التعلم فاننا نستطيع التحقق من هذا باعطاء اختبارات تحصيلية متكافئة لمجموعتين من التلاميذ يتشابهون فى كافة الخصائص ولكن مجموعة منهما تتعلم بالطريقة التعاونية بينما تتعلم الأخرى بطريقة مالوفة · ومن مقارنة نتائج الاختبارات التحصيلية للمجموعتين يمكن أن نختبر مدى صحة فرضنا · وهكذا فى سائر المجالات النفسية والتربوية ·

(ثالثا) دور الاختبارات فى الارشاد النفسى :

الارشاد النفسى Counseling هو العملية التى يمكن بها تنظيم أو عرض المعلومات الخاصة بالفرد على نحو يساعده فى الوصول الى حلول فعالة لمشكلات التكيف التى يعانى منها والتى تكون عادة فى المستوى السلوكى العادى أو السوى · وبهذا المعنى فان الارشاد النفسى جانب متخصص من جوانب الخدمات الفردية ، أو برامج التوجيه بوجه عام · أما اذا اتجهنا نحو مستوى الانحراف السلوكى الشديد فاننا ندخل فى باب علم النفس الطبى أو الطب النفسى والذى يتناول الأمراض النفسية والعقلية (كالعصاب والذهان) شديدة الخطورة سواء كانت ترجع لأسباب عضوية أو غير عضوية · واذا اتجهنا نحو المستوى السلوكى الأكثر سواء ندخل ميدان الخدمات الروتينية الخاصة بالتعريف المهنى واعطاء المعلومات عن المهن او الفرص التعليمية · ولكن المشكلات الأبعد مدى من مجرد اعطاء المعلومات والأقل من المرض النفسى أو العقلى أو سوء التكيف الاجتماعى (انحراف الأحداث أو الاجرام) قد تكون مشكلات ملحة ومتنوعة وغـــالبا ما تكون على درجة من الصـــعوبة والتعقيد ، والأغلب تكون ذات أثر طويل المدى · ومن هذه المشكلات ما يتعلق بالتخطيط للمهنة والتخطيط لنوع التعليم الملائم للفرد ، ومشكلات التكيف

الاجتماعى والانفعالى التى يواجهها المراهقون سواء فى بداية مرحلة المراهقة أو نهايتها ، والمشكلات الخاصة بصعوبات التعلم والتى تتطلب تعليما علاجيا من نوع معين ٠

فالمرشد النفسى الذى يتناول هذه المشكلات وأمثالها بالبحث والدراسة والتشخيص والعلاج هو واحد من مجموعة كبيرة من المتخصصين فى الخدمات الفردية منهم الأطباء والأطبــاء النفســيون والاخصــائيون الاجتماعيون والمتخصصون فى عيادات الأطفال ٠ والارشــاد النفسى مثله مثــل تلك التخصصات لا ينشط فى فراغ ولكن تحدده البيئة وظروف المجتمع التى تؤثر دون شك فى وسائل المرشد النفسى وأنشطته وأهدافه ٠ فالنظام التعليمى عموما يتضمن أنماط من المناهج الخاصة التى تحددها ضرورات اجتمــاعية أحيانا أخرى ٠ وتمثل الضرورة القانونية فى أن الطبيب لا يمكن أن يمارس هذه المهنة الا بعد التخرج فى كلية الطب ، على حين أن اصرار أصــحاب الأعمال على التخــرج فى أقســام الصحافة وادارة الأعمــال مثلا يمثل الضرورة الاجتماعية عندنا ٠ كما أن الأسر المصرية حين ترسل أبناءها الى المدرسة انما تعتبر عملية التعليم فى ذاتها صورة متخصصة للتأمين ضــد البطالة ٠ وبالتالى فان الآباء يلحون على أبنائهم أن يبــذلوا قصــارى جهدهم لمواصلة تعليمهم ٠

أما النظام التعليمى المصرى نفسه فانه يؤكد اساسا الدرجات المدرسية ومجموعها ونظام القبول حسب مكاتب التنسيق التى أصبحت من الثوابت الرئيسية فى نظامنا التربوى ، والتى امتدت من القبول بالمدارس الاعــدادية حتى القبول بالكليات والمعاهد العليا ٠

واذا كان لوظيفة الارشاد النفسى أن تحقق أغراضها فى مجتمعنا فلابد أن تعمل خلال المحددات التى تضعها هذه الاتجاهات والمطــالب والتقاليد الاجتماعية ٠ وسواء اتفقنا أو اختلفنا معها فانها تمثل الاطار المرجعى القائم الذى يعمل المرشد النفسى فى حدوده ، تماما كما يكون على الأطباء أن يتناولوا مشكلات الصحة الصناعية لصناعة معينة فى منطقة معينة ٠ هذا من حيث الشكل الذى يحدد خدمات الارشاد النفسى ٠ أما من حيث المضمون فان هــذه الخدمات وما يتعلق بها من تقويم وقياس تتأثر بأهداف المؤسسة (تربوية كانت

أم مهنية) وأغراضها • فاذا كان الهدف الرئيسى هو الانتقاء أو النصح أو التزكية أو الاستشارة ، فان معنى ذلك أن اجراءات الارشاد النفسى وعمليات التقويم والقياس المرتبطة بها تختلف عما لو كان الهدف الأكبر أكثر عمومية وأكثر اهتماما بنمو الأفراد وتحسين مستوياتهم •

والمرشد النفسى هو فى العادة اخصائى كلينيكى يتعامل مع فرد واحد فى المرة الواحدة ، ولذلك فهو عادة ما ينزع الى أن يكون أكثر اهتماما بجوانب السلوك الفردى ، ملتزما بمعايير متعددة الأبعاد لتكيف الفرد • وهو لا يهتم بالنجاح كما يقاس بالتحصيل المباشر فحسب ولكن بالنجاح كما يقاس بالرضا الفردى والبقاء والاستمرار فى المدرسة أو المهنة ، كما أنه لا يحاول أن يتنبأ بنتائج برنامج معين فى القراءة العلاجية مثلا فحسب ولكن بالتكيف الاجتماعى للفرد كما يتأثر بالخبرات الاجتماعية والنمو • ومن واجبه ألا يركز على الصراع العاطفى المباشر فحسب ، ولكن على آثار هذا الصراع فى النمو العاطفى والانفعالى بوجه عام •

ولكن ما هو الدور الذى يقوم به التقويم والقياس فى خدمات الارشاد النفسى ؟

يمكن القول بأن علاقة التقويم بالارشاد النفسى هى علاقة الوسيلة بالغاية ، أى المساعدة على التعرف على نواحى القوة والضعف سواء كانت داخل الفرد أو بين الأفراد • بعبارة أخرى يفيد التقويم والقياس فى جمع المعلومات عن الأفراد الذين يعانون من المشكلات الفردية الخاصة بالتكيف والتوجيه والنمو • كذلك مشكلات الجماعات من حيث التحصيل والانتقاء أو التوجيه المهنى والتربوى والتصنيف والتقسيم سواء فى الميادين التربوية أو المهنية أو العسكرية ، أى فى اتخاذ قرار ما • ولا يقتصر دور التقويم على مجرد جمع المعلومات وانما يمتد الى زيادة استبصار المرشد النفسى وفهمه لسلوك الأفراد ولسلوكه هو أيضا حتى يستطيع الوصول الى التشخيص المناسب • كما يسمح التقويم والقياس ، وخاصة اذا توفرت فيه شروط معينة سنناقشها فى هذا الكتاب ، بالوصول الى تنبؤ أدق مما لو استخدمنا أية وسائل لا تتوفر فيها تلك الشروط •

والواقع أن التشخيص والتنبؤ المعتمدين على الوقائع لها اهمية خاصة فى حياتنا الاجتماعية عموما · فمثلا لا يمكن الاعتماد على أى برامج للانتقاء أو التوجه الا اذا سمحت هذه البرامج بالتنبؤ الكمى الجيد ، وعـلم الطب يعتمد أساسا على التشخيص من ناحية وعلى القيمة التنبؤية للمقـــاييس الفسيولوجية من ناحية أخرى · كما أن كلا من الطب النفسى والخــــدمة الاجتماعية قد أفاد كثيرا من المعلومات الكمية التى نجمعها عن ســـلوك الأفراد ·

وهذا المنطق يصدق بالطبع على خدمات الارشاد النفسى · وهناك اتجاه يرى استخدام الاختبارات بأنواعها مسألة أساسية فى العلاج والارشـــاد النفسى ، ذلك لأنها أدوات يتحدد فى ضوئها تشخيص نوع الصعوبة الـذى يعانى منها العميل ، ومنها يستطيع المعالج أو المرشد النفسى أن يكون فكرة اقرب الى الدقة عن هذا العميل · كما أن الاختبارات ـ وخاصة اختبارات الشخصية ـ تساعد فى التعرف على مدى التقدم الذى يحرزه الشخص فى أثناء عملية الارشاد النفسى أو العلاج النفسى ·

ويمكن أن نلخص الأسباب التى من أجلها يؤكد أصحاب هذا الاتجاه أهمية الاختبارات فى عملية الارشاد النفسى فيما ذكره بوردين (٢١ : ٢٦٢ ـ ٢٧٩) وتايلور (٣٣) كما يلى :

١ ـ تمد الاختبارات المرشد النفسى بمعلومات معينة يحتاج اليها فى عمله مع الشخص وتساعده فى مهمته معه ·

٢ ـ تمد الشخص أو العميل بمعلومات معينة يحتاج اليها وخاصة فى اتخاذ قرارات معينة ·

٣ ـ قد تستثير الاختبارات العميل لاستكشاف ذاته ·

٤ ـ أن العميل قد يكتسب استبصارا وفهما لنفسه وهو يقوم بأداء الاختبار ·

الا أن هناك اتجاها آخر ينظر الى استعمال الاختبارات فى العلاج والارشاد النفسى نظرة مغايرة ، اذ يتردد أصحاب هذا الاتجاه فى الاعتماد كلية على منهج التقويم والقياس · وهذا التردد يعتمد على حجتين أساسيتين: أولاهما أن تأكيد المنطق الكمى والطبيعة العقلانية لعملية الوصول الى استبصار بموقف العميل قد يؤدى الى تضليلنا ، لأن ما نعتبره معلومات موضوعية محايدة هو فى حقيقةالأمر مرتبط بالكيان الانفعالى والعاطفى للعميل ، وثانيهما أن معرفتنا الراهنة بالشخصية وادوات تقديرها جميعا لازالت فى مراحلها الأولية المبكرة ·

وهذا الاتجاه يمثله فى الوقت الحاضر كارل روجز Carl Rogers (٢٨ ، ٢٩) وتلاميذه حيث يركزون الاهتمام على الجوانب العلاجية فى الارشاد النفسى ويتبعون طريقة خاصة من طرق العلاج تتميز بعدم التأكيد على الاختبارات لأنهم لا يقيمون وزنا كبيرا لعملية التشخيص والتصنيف فى العلاج · وجوهر طريقتهم هو أن الفرد مركز التقويم · أى أن التقويم يأتى من داخل الفرد ، وأن نموه وشفاءه أو حل المشكلة التى يعانى منها انما يأتى من قوى داخلية تنبعث منه عندما يوفر له المعالج الجو المطمئن الخالى من أى نوع من أنواع التهديد للذات أو التقويم للفرد · هذا الجو الذى يتحقق عندما يدعم المعالج عند العميل باستمرار وبعمق أنه يتقبله ويثق به ويهتم به باعتباره انسانا فى أزمة وجود أو أزمة حياة · ويساعده على التغلب عليها ليستمر نموه ·

هذا بايجاز شديد هو الجوهر الذى يدور حوله العلاج المركز حول العميل ، وفى ضوئه نفهم الأسباب التى جعلت الاختبارات عند أصحاب هذا الاتجاه لا تحتل مكانا هاما فى العلاج أو الارشاد النفسى ، ويمكن أن نلخص هذه الأسباب فيما يلى :

(١) المجال الذى ينشط فيه الفرد هو على درجة كبيرة من الاتساع وأن مواقف الأداء الفعلى عديدة ومتغيرة بحيث لا يمكن تمثيلها فى أى موقف اختبارى (الذى هو فى ذاته بحكم تعريفه ليس الا عينة من السلوك) ·

(٢) أن تأثير المتغيرات والعوامل التى يصعب أو يستحيل قياسها فى

الوقت الحاضر قد تكون على درجة كبيرة من القوة والأهمية ، مثل عوامل الحوافز والدوافع • الى حد أن الصفات السلوكية التى يمكن قياسها ـ اذ قورنت بما لا يمكن قياسه ـ قد تكون أقل أهمية بالنسبة للنشـــاط الواقعى فى للحياة (١٧) •

(٣) الاختبارات ليست الا عملية مقارنة بين العميل وآخرين فى جانب من الجوانب ، بينما المشكلة التى يعانى منها العميل هى فى حقيقتها مشكلة ذاتية نابعة من ادراكه الخاص بالمواقف والأشخاص ، ومن ادراكه لنفسه ، ومن المعانى التى يخلعها على هذه المدركات جميعا ، ولا يمكن ان نفهم العميل ومشكلته الا فى ضوء هذه الذاتية •

(٤) توحى الاختبارات الى العميل أن دور المعالج أو المرشد النفسى هو أن يقوم بعمل شىء له (أى للعميل) يحل مشكلاته وينهى معاناته ، ومن ثم يتخذ هذا العميل دورا سلبيا ينتظر فيه ما سيعمل له بدلا من أن يحس مسئوليته فى علاج نفسه متعاونا مع المعالج أو المرشد النفسى ، الذى لا يزيد دوره عن كونه ـ كما عبر روجرز ـ الأنا الأخرى للعميل • هذا الاتجـــاه السالب الذى تخلقه الاختبارات لدى العميل يجعله يتوقع « وصفة » لمشكلاته، ويصبح من أشق الأمور على المعالج أو المرشد النفسى نقل مسئولية الموقف واتخاذ القرارات الى العميل نفسه •

(٥) تحمل الاختبارات ايضا الى العميل شيئا من البعد بينه وبين المعالج أو تسبغ على العلاقة بينهما صفة الرسمية أو توحى بالعلو من جانب المعالج على العميل • وهذا النوع من العلاقة اذا جاز وجوده فى مجال الطب والأمراض الجسمية فانه غير مناسب فى علاقة المساعدة فى مجال الارشاد والعلاج النفسى •

(٦) قد يستعمل العميل الاختبارات عند أخذها وفى تفسيرها والتحدث حولها ستارا أو دفاعا أو هروبا من المشكلات الحقيقية التى ينبغى أن يواجهها مما يعطل سير عملية العلاج أو الارشاد النفسى •

بعض الوظائف التربوية للتقويم النفسى :

يؤدى التقـــويم النفسى وظائف عديدة فى مختلف مجالات الحيـــاة كالصناعة والادارة والتربية والعلاج النفسى وغيرها ، على نحو يدل على أهميته وقيمته للباحث النفسى فى هذه المجالات ونعرض فيما يلى بعض هذه الوظائف فى مجال بذاته وهو مجال التربية ، ويمكن للمهتمين بالمجـــالات الأخرى أن يجدوا بسهولة ويسر وظائف التقويم النفسى واضحة لهم ·

(أولا) دور التقويم فى زيادة الدافعية للتعلم :

يقوم التقويم بالوظائف الثلاث الرئيسية للدوافع فى التعلم وهى :

(أ) وظيفة التنشيط energizing اى زيادة المستوى العـــام للنشاط والجهد المبذول ، ويمكن التدليل على ما تحدثه الامتحانات السنوية والعامة من ازدياد النشاط والجهد فى فترات العمل الشديد التى يقضيها طلاب المدارس الثانوية والجامعات قبيل الامتحانات · وكثيرا ما تكون هذه الامتحانات فى معظم المدارس خاتمة المطاف ، ويكون للنجاح والفشل فيها اهمية كبيرة · كما تحدد هذه الامتحانات عموما متى يتعلم التلاميذ وماذا يتعلمون وكيف يتعلمونه ، ولذلك فان هذه الامتحانات اذا لم تقم بقيـــاس أهداف المنهج او المقرر قياسا حقيقيا وصادقا فاننا نشـــك كثيرا فى قيمة الدافعية التى تستثيرها ·

ويمكن القول من الوجهة المنطقية البحتة انه ما دام التقويم يزيد الجهد المبذول فان تكرار اجراءات التقويم يؤدى بالتالى الى زيادة الجهد الكلى للتلاميذ ، ولكن من المؤكد أن التكرار المناسب للتقويم يعتمد على طبيعـــة الأدوات التى تستخدمها ، وطرق التدريس ، وما اذا كان التلاميذ يصححون اختباراتهم أو اختبارات غيرهم ، والمادة الدراسية ، وقدرة التلاميذ بوجه عام · ولا شك أن اجراء التقويم بكثرة يقلل نسبيا من اهميته ، فالامتحانات اليومية مثلا ، قد لا يكون لها أى أثر منشط يزيد عن أثر الواجبات المدرسية اليومية · ويلخص كوك (٢٢) نتائج البحوث التى أجريت على أثر التقويم فى الدافعية فى أنه عندما تعطى كل أسبوع الاختبارات التحصيلية فى المواد

الدراسية فى مستوى الجامعة ثم تناقش النتائج وتستخرج الأخطاء الفردية، وعندما تتم صياغة أسئلة الامتحان النهائى على صورة تشبه تلك الاختبارات فان تحصيل الطلاب من ذوى المستوى المنخفض من القدرة يكون أعلى منه فى حالة التقويم الأقل تكرارا • أما الطلاب من ذوى المستوى المرتفع من القدرة فقد يتخلفون ما لم يكن فى كل اختبار أسئلة على درجة مناسبة من الصعوبة تتحدى قدرتهم ، كما أنهم يفيدون أكثر من المادة الاضافية التى يمكن للمعلم أن يدرسها اذا كرس وقتا أقصر للتقويم والعمل العلاجى بوجه عام• ولكن هل يمكن أن تعزى آثار التقويم المتكرر فى تحسين تحصيل التلاميذ الضعاف الى وظيفة التنشيط هذه ؟ الواقع أن معظم الباحثين يرون ان من المحتمل أن هذه الآثار ترجع الى توجيه التعلم واختيار الاستجابات الصحيحة (وهما الوظيفتان اللتان سنتحدث عنهما فيما بعد) أكثر من زيادة الطاقة والنشاط والجهد التى تستثيرها الامتحانات •

وقد تناول الباحثون بالدراسة ايضا أثر معرفة نتائج الامتحانات (والتقويم عموما) فى زيادة الدافعية للتعلم واكدت نتائجهم أن معرفة النتائج أفادت التعلم بمقدار النجاح الذى تتضمنه النتائج ، وقد افاد التلاميذ من ذوى المستويات المرتفعة اكثر من غيرهم أما أولئك الذين أحرزوا تقدما بطيئا فلم تستثرهم معرفة النتائج الا استثارة ضئيلة للغاية •

وقد درست هرلوك (فى ١١) أثر المدح العام أو اللوم العام فى التعلم• واكدت نتائجها أن المدح اكثر استثارة للتعلم من الذم ، وهو اشد تأثيرا فى التلميذ الأقل نضجا والأقل قدرة ، أما الذم فكان أكثر فائدة بصفة عامة مع التلميذ الناضج والقادر ، وقد لوحظ أن كلا من المدح والذم اكثر استثارة للتعلم من موقف الحياد أو التجاهل التام •

وخلاصة القول ، ان الأثر المنشط للتقويم يعتمد اساسا على درجة نجاح التلميذ ، فالذى يحصل على درجات عالية وبالتالى ينجح ويتقدم ويلقى المديح تكون استثارته ايسر من التلميذ الذى يحصل على درجات منخفضة وما تستتبع ذلك من فشل ورسوب وذم ولوم • وهذه النتيجة ذاتها تؤكد ضرورة وضع أهداف تربوية معقولة تتضمن على الأقل نجاحا جزئيا وممكنا للجميع •

(ب) وظيفة التوجيه directive ، وبها يتم توجيه سلوك المتعلم ونشاطه الى المسالك المرغوب فيها • فاذا كانت نتائج التقويم يمكن بها الحكم على المدارس والمعلمين والتلاميذ فان طبيعة أدوات التقويم تحدد الى حد كبير نوع العملية التربوية • فاذا كانت الاختبارات تتضمن أحكاما كيفية هامة فانها تفيد كثيرا فى تحديد الأهداف والطرق التربوية ، واذا كانت الاختبارات تعتمد على مضمون تقليدى بحيث تدور معظم أسئلتها حول الحقائق بغرض تحديد مقدار ما حفظه التلاميذ من مادة الكتاب المدرسى ، فان آثار هذا النوع من الاختبارات تختلف عن الاختبارات التى تقيس مهارات الدراسة وقدرات حل المشكلة ، لأن الأخيرة تؤكد المفاهيم وتطبيقاتها فى المواقف الجديدة وتوضح الأهداف الهامة وتركز الانتباه على الهدف الأقصى لعملية التربية وهو التعلم الدائم • ويمكن تلخيص الفرق بين نوعى الاختبارات فى أن النوع الأول يجمد المنهج المدرسى والنوع الثانى يذيبه •

ولا شك فى أن التقويم يؤثر فى اجراءات الدراسة وأهدافها كما سبق أن أوضحنا • ففى اعداد امتحانات من نوع المقال يكون على التلاميذ تحديد المادة وتنظيمها فى وحدات كبيرة وتأكيد العلاقات والاتجاهات والاستجابات الشخصية الخاصة بالتلميذ الفرد ، أما فى اعداد الاختبارات الموضوعية التى تعتمد على الحقائق فانها تتناول عينة من السلوك أشمل من اختبارات المقال ولكنها تتطلب من التلاميذ تأكيد التفاصيل والاسماء والتواريخ ونتائج التجارب • وقد لوحظ أيضا بعض الاختلاف فى طريقة اعداد التلاميذ لأنفسهم عند الاجابة على الأنواع المختلفة من الاختبارات الموضوعية ، فمثلا عند الاستعداد لاختبارات التكملة فان الطالب يحاول اتقان أسلوب الكلمة فى مقابل الكلمة ، أما فى حالة الاستعداد لاختبارات الصواب والخطأ أو الاختيار من متعدد بوجه عام فان التلميذ يهتم بالتعاريف والحقائق التفصيلية • ولا شك أن هذا من العوامل التى تحدث فروقا فيما نتعلمه وفى مقدار ما نحفظه لفترة من الزمن وفى القدرة على أداء أنماط مختلفة من الاختبارات ، وكلها ترتبط تجريبيا بنوع الاختبار الذى يستعد الطالب للاجابة عنه •

(ج) وظيفة الانتقاء selective ، وتعنى هذه الوظيفة تحديد الاستجابات التى سوف يتم تثبيتها وبقاؤها عند التلاميذ والاستجابات التى

سوف تحذف ، وترتبط هذه الوظيفة بالجانب التشخيصى للتقــويم ، الا أن الفرق بين الانتقاء والتشخيص هو أن الاختبار التشخيصى يركز الاهتمام على تشخيص الصعوبات كما يقترحها المعلم أو الاخصائى النفسى المدرسى ، أما ما نتحدث عنه الآن فهو أن التلاميذ والمعلم والاخصائى النفسى جميعا يمكنهم الحصول على مقدار كبير من المعرفة من اخطاء التلاميذ (واسبابها اذا كان هذا ممكنا) فى موقف الاختبار . وينتج أكبر مقدار من التعلم من موقف الاختبار ذاته حينما يسمح للتلاميذ بتصــحيح أوراق اجاباتهم ثم يلى ذلك مناقشة الأخطاء ثم العمل العلاجى . وتؤكد البحوث التجريبية انه كلما كانت معرفة التلاميذ لجوانب الصواب والخطأ فى استجاباتهم وأسبابها معرفة مباشرة فى موقف الاختبار فان ذلك يؤدى الى تثبيت الاستجابات الصحيحة أو السلوك المرغوب فيه وحذف الأخطاء واستبعادها .

(ثانيا) دور التقويم فى تصميم المواقف التعليمية :

يعد مبدأ الفروق الفردية من المبادىء الأساسية فى علم النفس الحديث ، وهو يعنى من الوجهة التربوية التباين فى جماعات الاعمار الزمنية والفرق الدراسية فى الذكاء والقدرات والتحصيل والاتجــاهات والميول وســمات الشخصية المختلفة . وأى رأى يعتبر محض الفرقة الدراسية مرحلة لهــا خصائصها المحددة التى تنطبق على جميع الأفراد المنتسبين لها رأى خاطىء ويعطل جهود المؤسسة التربوية وسعيها لاشباع حاجات التلاميذ كأفراد . وأى افتراض حول وظيفة المعلم على أنه متخصص فى فرقة دراسية معينة انما يلغى امكانية نجاحه فى التعامل مع التلاميذ الذين قد يشذون عن المستوى الافتراضى لهذه الفرقة . وأى جهد من جانب السلطات التعليمية لجعل المقرر الدراسى جامدا ثابتا وعاما على جميع تلاميذ الفرقة الدراسية لا يضع فى الاعتبار أن هؤلاء التلاميذ ليسوا جميعا على نفس الدرجة من القدرة على أداء ما يطلب منهم بنجاح .

وقد أجريت عدة محاولات لحل مشكلة الفروق الفردية فى مواقف التعلم نذكر ما يلى :

١ ـ **مراعاة القدرة العقلية العامة :** من الافتراضــــات التى وجهت

الممارسات التربوية فى الماضى أن الجماعات المدرسية تكون اكثر تجانسا فى التحصيل اذا تجانست فى القدرة العامة ولذلك كانت تستخدم طرق تربوية مقننة مثل توحيد المقررات والامتحانات بحيث تتفق مع هذا المستوى العام ويتضمن هذا الافتراض مسالتين أولاهما أن التباين داخل الفرد لا يعتد به ، وثانيهما أن جميع الخصائص السلوكية التى تهتم بها المدرسة ترتبط فيما بينها ارتباطا عاليا وتنتظم حول عامل يحدد المستوى العام لكفاءة الفرد ٠ والواقع أن الأدلة العلمية الحديثة تدحض هذا الافتراض وتضميناته ٠ وقد جاءت هذه الأدلة من مصادر عدة أشهرها البحوث الحديثة فى القدرات العقلية التى أكدت أن الاعتماد على القدرة العقلية العامة وحدها لا ينتج مجموعات متجانسة تجانسا كاملا (١٠) ٠

٢ ـ **مراعاة مستويات التحصيل السابقة** : من ناحية اخرى فان المعلمين ورجال الادارة التعليمية والمسؤلين التربويين يرجعون التفاوت الكبير فى مدى القدرة لدى المتعلمين الى عدم الدقة فى اجـــراءات نقلهم من المســـتويات التعليمية وهم فى هذا الصدد يطالبون بأن تكون للفرقة الدراسية الواحدة مستويات محددة يجب أن يصل اليها التلاميذ قبل انتقالهم الى فرقة أعلى ، وعلى الرغم من جودة هذه الفكرة الا انها صبت احيانا فى القالب التقليدى حيث يحدد المقرر المدرسى ما يجب أن يتعلمه الجميع فى كل مادة فى كل فرقة دراسية وأى الكتب المدرسية تستخدم ، وأى صفحات أو فصول يجب ان تعطى فى وقت محدد من السنة ، وتكون محصلة الأمر ان يتعلم التلميذ المعـارف المتضمنة فى هذه الكتب مثل : تهجى الكلمـــات ، المســـائل النموذجية فى الحساب ، الأسماء التواريخ ، أسباب الحروب ونتائجها ، الدول وعواصمها، والمدن وخصائصها الخ ٠

والسؤال الآن هو : هل توجد بالفعل مستويات للتحصيل لا يتباين فيها التلاميذ بحيث تؤدى الى تحقيق التجانس داخل الجماعات المدرسية ؟ للاجابة على هذا السؤال يمكن صياغة القاعدة العامة الآتية التى نســتخلصها من نتائج البحوث العديدة التى أجريت فى هذا الميدان ٠

« اذا كانت الاستجابات التى نتعلمها بسيطة والأهداف التى ننشدها محددة بحيث يمكن لأكبر عدد من التلاميذ أن يصلوا اليها خلال فترة التعلم

فان تباين المجموعة يتناقص ، أما اذا كان العمل على درجة من التعقد والتركيب والأهداف غير محددة بحيث لا يمكن أن يصل اليها الا التلاميذ الأقدر فان تباين المجموعة يتزايد » ٠

ومعنى ذلك ان الرأى القائل بان التلامذ يمكنهم ان يصبحوا اكثر تجانسا اذا استخدمنا أساليب التدريس التى تؤدى بهم الى مستوى واحد مقنن انما يفترض محدودية الأهداف وضيق نطاق التدريس فحتى يمكن أن تكون المهام فى متناول التلاميذ الأقل قدرة يلجأ المعلم الى اسـاليب التعلم الروتينى ، كما أن الكتب المدرسية تكون فى العادة من مستوى أدنى ، والمواد التى يتم اختيارها للمنهج تكون من النوع الذى يمكن أن يتقنه حوالى ٨٠٪ من التلاميذ ، ويتم تدريب التلاميذ على الأسئلة واجاباتها ، مثـل اسئلة الامتحانات ، وعادة ما يكون معظم وقت المعلم مكرسا للتلميذ بطىء التعلم وفى هذه الحالة عندما يقاس تقدم التلميذ فى فترة زمنية معينة فى ضوء الأهداف المحدودة فان التلاميذ ذوى الدرجات المنخفضة فى البداية يحرزون اعلى تقدم ممكن ومعنى ذلك أن هناك علاقة سالبة بين الدرجات المبدئية وبين درجات الاستفادة او الكسب فى نهاية التعلم ٠ كما أن العلاقة بين الاستفادة فى هذه الحالة والذكاء علاقة منخفضة وكثيرا ما تكون سالبة ٠ وهذه النتائج تؤكد أن مثل هذه الممارسة لا تستثير التلاميذ من ذوى القدرة العاليـة او المتوسطة الا قليلا ان لم تهملهم اهمالا تاما ٠

وقد تكون اخطر النتائج المترتبة على ذلك أن ما يتعلمه المتعلمون قليل القيمة ، ضعيف الأثر ، ضيقا فى مدى الاحتفاظ به وتستخدم فيها اسـئلة الحقائق والتفاصيل والمعلومات المحدودة التى تدور حول أذكر ، وصف ، وعرف ، ومن ، وماذا ، ومتى ، وأين ٠ ولا يستخدم التلاميذ فى الاجابة عليها الا الذاكرة ٠ وتكون آثار التعلم فى هذه الحالة مؤقتة وخاصة مع استخدام الذاكرة الصماء ٠ لأن ما ينساه التلميذ فى هذه الحالة تكون ما بين ٤٠٪ خلال ثلاثة اشهر و ٨٠٪ خلال ثلاث سنوات ، ويسمى باكنجهام (فى ١٠) ذلك « اعظم فقد تربوى » ٠ ومنحنيات النسـيان لمعظـم المواد الدراسية فى هذه الحالة تقترب كثيرا من منحنيات نسيان المواد عديمة المعنى ٠

أما اذا كان التعلم يتضمن عمليات حل المشكلة وغيرها من العمليات

العقلية العليا فان آثاره تكون أكثر دواما • أى أنه ما لم يعمل التعلم على تمايز وتكامل الاستجابات القديمة والجديدة فى نمط سلوكى منظم لا تكون لهذا التعلم قيمة أو لاستمرار • وهذه النتائج تهمنا كثيرا فى الوقت الحاضر مع انتشار مفهوم « التعلم للاتقان » وحتى لا يتحول هذا الشعار النبيل الى محض وصول الى مستوى متواضع من القدرة •

خلاصة القول أن حل مشكلة الفروق الفردية فى مواقف التعلم لا يكون بتحقيق تجانس التلاميذ فى القدرة العامة أو باستخدام الطرق المقننة وانما بالاجابة على سؤال أهم هو كيف تم التعلم وكيف صممت مواقفه ؟ فاذا سادت التعلم العمليات العقلية العليا وكانت أهدافه من العمومية والاتساع والارتفاع أدى ذلك الى تحقيق نتائج أكثر استمرارا ودواما والى استثارة المتعلمين لبذل قصارى جهدهم فى التعلم فاننا بذلك نصل بالمتعلمين الى الحد الذى يظهر قدراته وتفرده وعلى نحو فيه التباين الابداعى أكثر مما فيه من التجانس الجامد •

(ثالثا) دور التقويم فى تصنيف المتعلمين :

لجأ بعض الباحثين فى حلهم لمشكلة الفروق الفردية الى افتراض أن التباين داخل الفرد والتباين بين الأفراد يدعوان الى ضرورة مقابلة الحاجات الفردية • ويعنى ذلك أن نقبل حقيقة وجود التباين الواسع فى السلوك الانسانى •

ولمواجهة هذا المطلب فى المدرسة الابتدائية مثلا اهتم الباحثون بخصائص نمو التلاميذ من ناحية ، وتكيفهم المناسب للثقافة التى يتأثرون بها من ناحية أخرى • ففى عملية التفاعل الاجتماعى العام خلال سنوات التكوين تكون السمات الأكثر أهمية فى تحديد عضوية الطفل ومكانته فى الجماعةهى للعمر الزمنى والنمو الجسمى العام ، واذا أضفنا الى ذلك النمو الاجتماعى العام ، تكون هذه السمات الثلاث فى رأى المهتمين أساس التصنيف والتجميع التربوى • الا أن ما يجب ملاحظته أن العمر الزمنى لا يرتبط ارتباطا كاملا بالنمو الجسمى والاجتماعى للطفل • ولهذا تعطى الأولوية لمستوى النمو

لا العمر الزمنى وبخاصة النمو الجسمى والاجتماعى ، حتى يمكن للطفل أن يعيش ويعمل فى مجموعة ينتمى اليها تقبله ويقبلها ٠

ويضاف الى عامل النمو (وخاصة النمو الجسمى والاجتماعى) فى التصنيف باعتباره العامل الأولى والاساسى عامل آخر هو مستوى التحصيل النوعى فى المجالات المختلفة تبعا لقدرات التلاميذ فى مختلف المواد الدراسية٠ ويجب أن نتذكر دائما أنه حتى لو تم التجميع والتصنيف فى المدرسة الابتدائية بهذه الطريقة فلابد أن يتركز اهتمامنا على التلميذ الفـــرد ٠ وبالطبع فان ملاحظات المعلم لتلاميذه وحكمه عليهم يكون لها دائما اهمية قصوى فى موقف التعلم ٠

دور الحاسب الألكترونى فى التقويم النفسى :

يتزايد فى الوقت الحاضر استخدام الحاسب الالكترونى (الكمبيوتر) فى المجالات المختلفة لعلم النفس سواء داخل المعمل أو خارجه ٠ ولعل أكثر هذه المجالات استخداما للحاسب الالكترونى مجال التقـــويم النفسى ، ومن المتوقع أن يزداد هذا الاستخدام شيوعا وانتشارا فى المستقبل ٠

وتمتد الاستخدامات الراهنة للحاسب الالكترونى على نطاق عمليـــة التقويم النفسى كلها ، الا أن أكثر هذه الاستخدامات وضوحا فى عملية تصحيح الاختبارات وتحليل نتائجها ٠ وتتوافر لدى ناشرى الاختبـــارات فى الوقت الحاضر امكانات تزويد الباحثين بنتائج الاختبارات على أى صورة يشاءون ابتداء من نتائج المؤسسات وحتى نتائج الأفراد ، وذلك بعد تصحيح أوراق الاجابة التى عادة ما تتخذ صورة معينة تصلح للمعالجة بواسطة الكمبيوتر٠ ويستخدم الحاسب الالكترونى ايضا بسهولة ويسر فى عمليات تحليل اسئلة الاختبار وانتقاء الملائم منها كما يمكن المقارنة بين المعايير المحلية والمعايير القومية للاختبار ، ويستطيع الحاسب الالكترونى أيضـــا أن يزود المدارس والمؤسسات المختلفة بنتائج الاختبار لتقوم بدورها فى التغذية الراجعة ، وقد يسر ذلك على المعلمين والاخصائيين أن يستثمروا هذه النتائج فى تحســين برامج التعليم والتدريب ٠

ويتزايد فى السنوات الأخيرة استخدام الحاسب الالكترونى فى عمليات اعداد الاختبار ، وخاصة الاختبارات التحصيلية ، واستخدامها داخل المؤسسات نفسها (كالمدارس) ومع ظهور الحاسبات الالكترونية المصغرة تتوافر فى الوقت الحاضر اعداد منها فى معظم المدارس الأوربية والأمريكية، تفيد المعلمين فى سرعة تصحيح الاختبارات التى يقومون باعدادها ، ناهيك عن الحاسبات الالكترونية الكبيرة التى تستخدم فى تصحيح وتحليل نتائج الاختبارات ابتداء من المستويات المحلية وحتى المستوى القومى العام ·

وظهرت فى السنوات الأخيرة ايضا نظم اكثر تقدما تشمل ما يسمى « بنوك الأسئلة » او « مخازن مفردات الاختبارات » ، والتى تقوم فيها الحاسبات الالكترونية بالدور الحاسم ، وتفيد هذه «البنوك » او « المخازن » فى انتقاء المفردات أو الأسئلة لاعداد اختبار بمواصفات معينة (مثل أهداف تعليمية معينة ، محتوى معين ، مستوى صعوبة معين ، الخ) · وعادة ما يشمل النظام الكامل لبناء الاختبارات عن طريق الحاسب الالكترونى المكونات الآتية :

١ ـ بنك اسئلة ·
٢ ـ نظام لتوليد الأسئلة ·
٣ ـ بنك بخصائص الأسئلة ·
٤ ـ نظام لانتقاء الأسئلة ·

ومن المجالات الهامة التى امتد اليها استخدام الحاسب الالكترونى فى التقويم النفسى مجال تطبيق الاختبارات ذاته · فمن الملاحظ أن عددا كبيرا من المدارس فى الولايات المتحدة واليابان وبعض الدول الأوربية المتقدمة تطبق الاختبارات المعتادة عن طريق الكومبيوتر وخاصة البرامج التعليمية التفريدية التى يديرها الحاسب الالكترونى (أى حين يستخدم الكومبيوتر كمعلم) · ولعل مما هو اكثر اهمية استخدام الكومبوتر فى تطبيق ما يسمى الاختبارات « التفصيل » tailored او الاختبارات التكيفية التتابعية ، فالمرونة الفائقة التى عليها الحاسب الالكترونى تسمح له بأن يطبق على كل مفحوص الأسئلة او المفردات الاكثر ملاءمة لمستوى قدرته · وفى هذا الحالة يعتمد الكومبيوتر على استجابة المفحوص على سؤال معين ، ويتحدد فى ضوء هذه

الاستجابة نوع السؤال أو مجموعة الأسئلة التى تتلوه · ومثل هذه النوع من الاختبارات « التفصيل » توفر الوقت والجهد اللذين يبذلان فى الاجابة على اختبارات شديدة السهولة أو الصعوبة بالنسبة للمفحوص ·

ويوجد استخدام اخر للحاسب الالكترونى يعد اكثر حداثة وهو تزويد الفاحصين ببعض التفسيرات المفيدة فى الارشاد النفسى ومن ذلك ان يقدم الكومبيوتر تقريرا لفظيا عن أداء المفحوص يصـــــف المهن الاكثر ارتباطا وملاءمة لاستعداداته وميوله وتطلعاته التربوية يســـــمى « تقرير التخطيط المهنى » · وقد لوحظ أن الكومبيوتر يستطيع القيام بعمليات المقارنة بين انماط درجات الاختبارات المختلفة للمفحوص ويقوم بالربط بينها وبين المتطلبات المهنية المختلفة على نحو أيسر وأسرع ، بل واكفأ ، من المرشد النفسى · ومن المتوقع أن يوظف فى المستقبل القريب هذا الجانب من استخدامات الكومبيوتر فى التربية وخاصة ما يتصل بتشخيص صعوبات التعلم المدرسى · وتوجد بوادر لذلك فى الوقت الحاضر بالنسبة لاختبارات المهارات الاساسية ·

بعض المشكلات المعاصرة فى ميدان التقويم النفسى :

يواجه ميدان التقويم النفسى فى الوقت الحاضر عددا من المشـــكلات تتصل فى جوهرها باستخدام الاختبارات ، ويمكن ان نلخص هذه المشكلات فيما يلى :

١ ـ مستوى جودة الاختبارات :

شهد مطلع الستينيات بداية حملة منظمة ضد الاختبارات المقننة سرعان ما اشتدت وزاد سعيرها فى السبعينات ثم الثمانينات ، وقد وجهت سهام النقد على وجه الخصوص الى الأسئلة الموضوعية من نوع الاختبار من متعدد التى تتألف منها معظم هذه الاختبارات ، واتهمت بأنها تعاقب المفحوص الاكثر ذكاء وابتكارية · وقد دعم النقاد هذا الرأى بما يلاحظ من أن هذا المفحوص الذكى المبتكر يظهر فى اجابته على أسئلة الاختبارات المقننة اختلافا عن الاجابات التى تتحدد صحتها مقدما والتى تتكون منها مفاتيح هذه الاختبارات ، وعادة ما يحكم على هذا النوع من الاجابات التى لا تتفق مع المفتاح بالخطأ بالرغم

من أننا لو حللناها بعمق وأناة نجد فيها قدرا كبيرا من الصواب • وهذا النقد بالطبع له ما يبرره وخاصة أنه وجدت فى بعض الاختبارات المقننة ذات الشهرة الكلاسيكية أسئلة من هذا النوع • ناهيك بالطبع عما يمكن أن تجده فيما يتوافر لدينا •

الا أن هذا النقد لا يقلل من شأن الاختبارات الجيدة الصنع المتقنة البناء ، وقد كانت له تضميناته الايجابية ، فأصبح ناشرو الاختبارات فى وقتنا الحاضر يقدمون لمستخدميها بيانات تحليلية تفصيلية عن الأسئلة التى يتألف منها الاختبار مع تحليل احصائى ومنطقى لها • كما أن هذا النقد جعل مستخدمى الاختبارات أكثر حرصا وحذرا فى استخدامها بالاضافة الى شعورهم بالحاجة الى عدم الاعتماد على نتائج الاختبارات وحدها ، وانما لابد من تدعيمها بمعلومات اضافية من المصادر الأخرى •

ويرتبط بهذا النقد نقد آخر يتصل بمدى شمول الاختبارات لنطاق واسع من السلوك الانسانى • اننا نعلم أن الاختبارات لا تقيس فى معظم الحالات الا عينات من السلوك الانسانى (اذا تجاوزنا الاختبارات المرجعة الى المحك ذات الاستخدامات الخاصة فى مجال التعليم والتدريب) ، ومع ذلك نقع فى الخطأ الشائع الذى يتمثل فى أننا نتوقع منها ما لم تقصد اليه • ولعلنا نكون أكثر علمية ودقة اذا استخدمنا الاختبارات فى تحقيق الوظائف المحددة التى تسعى اليها ، انها حينئذ تصبح أدوات جديدة للعلم • ولعل هذا القول يعود بالنقاد الذين يوجهون الى اختبارات الاستعدادات ومنها اختبارات الذكاء فى وقتنا الحاضر عاصفة من الهجوم الى جادة الصواب • ان اختبار الاستعداد الذى يستخدم فى أغراض الانتقاء لمؤسسة تعليمية أو مهنية مثلا لا يحقق الغرض منه اذا افترض فيه كل من الناقد والمدافع جميعا أنه يقيس جميع المتطلبات اللازمة للنجاح فى دراسة معينة أو مهنة معينة ، فهذا القول يتجاوز المعطيات والبيانات التى توفرها الاختبارات • وحتى لا نقع فى أى تطرف جديد علينا دائما أن نتذكر أن الدرجات التى نحصل عليها من الاختبارات ليست الا نوعا من المعلومات – من بين أنواع عديدة أخرى متاحة للاخصائى النفسى ، ولا يجب عليه أن يصدر حكمه أو يتخذ قراره بناء على نتائج الاختبارات وحدها •

أن الاختبارات النفسية ـ كغيرها من وسائل أدوات البحث ـ قد تتفاوت في درجة الدقة ، وعلينا الا نتوقع منها ما ليس من طبيعتها ، وعلينا ان نجعل نصب أعيننا دائما أنه اذا كانت المقاييس الفيزيائية ـ وما هى عليه من سمعة وتاريخ ـ تعانى من أخطاء القياس وأخطاء الملاحظة وأخطاء التنبؤ ، فما بالك بالاختبارات النفسية وهى مقاييس للسلوك الانسانى بكل ما فيه من تعقد وتغير ورقى ٠

٢ ـ اثر الاختبارات في المفحوص :

يرجه النقاد الى الاختبارات نقدا آخر يتلخص فيما تحدثه هذه الاختبارات من آثار غير مرغوب فيها ـ ان لم تكن ضارة ـ فى المفحوصين من مختلف المستويات العمرية ٠ ولعل أشهر هذه الآثار ذكرا هو القلق ٠ وبالطبع لا ينكر أحد أن الموقف الاختبارى يحدث فى المفحوص قدرا من القلق، الا أنه من الحقائق السيكولوجية الهامة أن قدرا معقولا من القلق يدفع المفحوصين الى مزيد من التحسن فى الأداء ٠ وقد يكون الأثر الضار للقلق مرتبطا بالمفحوصين ذوى المستويات العليا فيه حيث يتداخل القلق الزائد مع الأداء الاختبارى ٠ وقد لا يكون فى ذلك مسئولية مباشرة للاختبار ذاته ، فكل ما يفعله القلق مع هؤلاء المفحوصين ذوى القلق المرتفع ان يضيف اليه المزيد ٠ ويوجد فى فن تطبيق الاختبارات قواعد أساسية لو التزم بها الفاحصون يمكنهم خفض القلق الاختبارى ،ويمكن أن يتم ذلك فى جميع مراحل تطبيق الاختبار ، سواء عند التهيؤ له ، أو التدريب على الأمثلة التوضيحية أو استخدام الوقت الحر فى الأداء الاختبارى (اختبارات القوة مثلا) ٠ ولحسن الحظ يتحول ناشرو الاختبارات فى الوقت الحاضر الى تزويد المفحوصين باختبارات يمكنهم التدريب عليها وتتوافر فى مكتبات الجامعات الأمريكية والأوربية أمثلة كثيرة من هذه الاختبارات التدريبية ، وبالاضافة الى ذلك فان مؤلفى اختبارات الأداء الاقصى يتحولون أيضا فى الوقت الحاضر الى التركيز على اختبارات القوة (التى لا تتحدد بوقت معين) بدلا من اختبارات السرعة ٠ وفى جميع الأحوال لابد للفاحص من ملاحظة المفحوص بعمق أثناء الموقف الاختبارى ، واعادة النظر فى درجات المفحوصين ذوى الانفعالية العالية وخاصة ان قلق الاختبار ليس له الا اثر مؤقت فى الصحة

النفسية للمفحوص • ومرة أخرى فان المسألة تعتمد على الاستخدام الجيد والتوظيف الرشيد للاختبار النفسي ونتائجه •

٣ - الأثر الاجتماعي والثقافي للاختبارات النفسية :

لعل هذه المسائل هي أخطر المسائل على الاطلاق ، وتأتي اهميتها من أنها شغلت الباحثين في مجال التقويم النفسي وخارجه سنوات طويلة تكاد تتساوى مع العمر الزمني لحركة القياس النفسي ذاتها • فكثيرا ما يقال ان الاختبارات النفسية غير عادلة في تعاملها مع الفئات الاجتماعية داخـــل المجتمع ومع السلالات البشرية بين المجتمعات المختلفة • وقد تزايدت هذه الانتقادات في السنوات الأخيرة في بعض المجتمعات التي مورست فيها الاختبارات النفسية لسنوات طويلة ، وعلى الأخص بريطانيا والولايات المتحدة الأمريكية ، ويرجع ذلك الى زيادة المطالبة بالحقوق المدنية وتهيئة تكافؤ الفرص للجميع •

وتتلخص الاتهامات الاجتماعية التي يوجها النقـــاد الى الاختبارات النفسية بأنها متحيزة اجتماعيا وتمثل حواجز ضد الفرص التعليمية والمهنية لعدد من ابناء الفئات المحرومة داخل المجتمع الواحد •

وحتى يمكن لنا مناقشة هذه الاتهامات بشيء من الموضوعية نذكر أن مسألة التحيز أو عدم العدالة في الاختبار تشير الى مســـألتين مختلفتين ، أولاهما تتصل ببناء الاختبار ذاته حيث يتألف من أسئلة أو مفردات تلائم فئات اجتماعية دون سواها ، وثانيتهما تتصل بالاســـتخدام الظالم لنتائج الاختبارات وخاصة فيما يتصل باصدار الأحكام أو اتخاذ القرارات ، وهي مسألة تخرج بالباحث من نطاق العلم الى التميز الصريح •

أما فيما يخص المسألة الأولى فاننا نجد أن معظم ما يكتب عن تحيز محتوى الاختبار يتركز على أداء بعض الجماعات أو الأفراد في الاختبــار يكون متخلفا ، فيحصــــلون على درجات منخفضـــــة بالمقارنة بغيرهم من الجماعات أو الافراد • الا أن هذا ليس حجة ويتجاوز الدليل الموضوعي ، فالأداء المنخفض في ذاته ليس برهانا كافيا على تحيز الاختبار ، والا كان كل اختبار في الحساب يصبح متحيزا ضد الضعاف في الرياضيات ، وكل

اختبار فى المفردات اللغوية يصبح متحيزا ضد أصحاب المحصول اللغوى المتواضع • ان هذا القول لا يصبح صحيحا الا اذا تدخلت الأفراد فى الاختبار عوامل أخرى لم يتحكم فيها الباحث ، كأن تصاغ المشكلات الحسابية فى قالب لغوى يحتاج الى قدرة لغوية عالية ، أما اذا تحرر الاختبار الحسابى من هذه العوامل (أى كان صادقا بالمعنى الدقيق للكلمة) فلماذا لا يكون ضعف الأداء فيه دليلا على نقص فى القدرات والمهارات الحسابية تحتاج الى نوع من التربية التعويضية أو التعلم العلاجى •

الا أن ما نحب أن ننبه اليه هو أن بعض الاختبارات المقننة تركز على المحتوى والقيم الثقافية المألوفة والشائعة لدى الانسان الغربى الأبيض من الطبقة المتوسطة (فقد كانت الاختبارات منذ نشأتها الأولى غربية المنشأ قالب لغوى يحتاج الى قدرة لغوية عالية ، أما اذا تحرر الاختبار الحسابى والسلالات البشرية الأخرى • ولذلك لابد من أن يتنبه الباحثون فى الثقافات الأخرى الى عدم اللجوء الى « النقل الأعمى » أو « الترجمة الحرفية » لهذه الاختبارات ، والا أوقعوا أنفسهم وثقافتهم فى مأزق اجتماعى خطير • ولعلنا بهذا نشير الى بعض الممارسات الخاطئة فى حركة الاختبارات النفسية فى الثقافتين المصرية والعربية حيث ينطبق على معظمها خصائص النقل الأعمى والترجمة الحرفية • بل اننا بهذا ندعو الى ابداع ثقافى جديد فى مجال الاختبارات النفسية ، فليس كل اختبار ظهر فى عاصمة أوربية أو فى جامعة امريكية قابل للنقل ، فبعضه وراءه من الافتراضات النظرية فى علم النفس والأسس الثقافية للمجتمع ما يجعل نقله ـ حتى ولو كان مع بعض التعديل ـ نوعا من « المجزرة » العقلية كما قال ـ بحق ـ عالم النفس التركى جوندوز فاسف فى بحثه عن الاختبارات النفسية فى العالم الثالث والذى ألقاه فى المؤتمر الدولى الثالث والعشرين لعلم النفس الذى انعقد بالمكسيك عام ١٩٨٤ •

الفصـــل الثـــانى

مدخل الى الاحصاء النفسى

أشرنا فى الفصل السابق الى أن القياس النفسى يعتمد فى جوهره على البيانات الكمية التى تتوافر عن السلوك الانسانى ، وهى ، كالبيانات الكيفية يجب أن تتاح للاخصائى النفسى كى يصدر أحكامه على هذا السلوك ويتخذ القرارات العملية حوله ٠

وبالطبع فان البيانات الكمية التى تستند اليها عملية التقويم النفسى قد تكون أدق من البيانات الكيفية ، ويتوقف ذلك على طبيعة الظاهرة النفسية موضع البحث والدراسة ، ويجب على الأخصائى الا يتعسف فيفرض على هذه الظاهرة طبيعة غير طبيعتها ، ومن ذلك أننا حين نقارن بين شخصين فى « طول القامة » فان من الأفضل أن نصف الشخص الأول مثلا بأن طوله ١٦٠ سم ونصف الشخص الاخر بأن طوله ١٨٠ بدلا من تقول ان الشخص الثانى أطول من الأول ، فالعبارة الأولى استخدمت لغة الكم ، وهى أقرب الى المقام والى طبيعة الظاهرة المقيسة ، ولكننا عند المقارنة من شخصين فى القدرة على القيادة ، يكون وصفنا اللفظى المفصل المعتمد على بصيرة الباحث أجدى وأفضل ٠

وتحتاج البيانات التى يحصل عليها الباحث الى تحليل ، ويستخدم علم الاحصاء فى تحليل البيانات الكمية خاصة أما البيانات الكيفية فلها طرقها الخاصة ٠

ونقدم فى هذا الفصل عرضا موجزا لطرق تحليل البيانات الكميـة مصنفة تبعا لمستوى المقياس المستخدم ٠ وقد سبق أن أشرنا الى أن مستويات القياس هذه أربعة وهى المقاييس الاسمية ، ومقاييس الرتبة ومقاييس المسافة ومقاييس النسبة وسوف نقتصر على الطرق المستخدمة فى الاحصاء الوصفى خاصة كما يتمثل فى الاتجاهات الرئيسية الثلاثة ، النزعة المركزية والتشتت والعلاقة، مع ملاحظة ان الطرق الاحصائية التى تصلح للمستويات الدنيا تصلح أيضا للمستويات العليا والعكس غير صحيح ٠

تحليل بيانات المقاييس الاسمية

أشرنا الى أن المقاييس الاسمية تستخدم الأعداد لتشير الى أفراد أو فئات ، وفى حالة استخدام الفئات ، فاننا نقوم « بعد » عدد الحالات التى تقع فى كل فئة ، ويطلق على هذا العدد « التكرار » وتوجد طرق احصائية عديدة للتعامل مع التكرارات المصنفة الى فئات ، اما الاعداد التى تشير الى الفئات ذاتها فلا تخضع لأى تحليل احصائى لأنها فى الواقع محض رموز وقد يشير المقياس الاسمى الى ما يسمى القيم المنفصلة وفى هذه الحالة ينطبق على أسئلة الاختبار اذا كانت الاجابة عليها من نوع « الكل أو لا شيء » مثل « اجابة صحيحة أو خاطئة » أو « نعم أولا » وعوملت منفصلة ، أى كل سؤال على حدة ٠

ويحتاج الباحث كما أشرنا فى تحليله لبيانات المقاييس الى استخدام الطرق الاحصائية فى تحديد الاتجاهات الرئيسية الثلاثة : النزعة المركزية ، والتشتت ، والعلاقة ٠ وقد لا تتضح هذه الاتجاهات جميعا فى المقاييس الاسمية ، بينما نجدها كلها فى مقاييس الرتبة ، ومقاييس المسافة والنسبة كما سوف تلاحظ ٠

النسب والنسب المئوية :

عادة ما يستخدم فى تحليل التكرارات النسب والنسب المئوية ، والمثال التالى يوضح ذلك :

لاحظ احد المدرسين اختلاف عدد التلاميذ والتلميذات فى فصلين من فصوله ، كما يتبين من الجدول التالى :

	تكرار التلاميذ	تكرار التلميذات	المجموع
الفصل (أ)	٣٠	٣٠	٦٠
الفصل (ب)	٣٠	٢٠	٥٠

انه لا يستطيع أن يستنتج من الجدول السابق ان الوضع النسبى للذكور فى الفصلين متساو مادام عدد التلاميذ الذكور متساو ، ويفيد فى المقارنة ان يقوم بتحويل التكرارات المتضمنة فى الجدول السابق الى نسب او نسب مئوية كما يلى :

	نسب التلاميذ	نسب التلميذات	النسبة المئوية للتلاميذ	النسبة المئوية للتلميذات
الفصل (أ)	٥ر	٥ر	٥٠٪	٥٠٪
الفصل (ب)	٦ر	٤ر	٦٠٪	٤٠٪

وقد حصلنا على نسبة الذكور مثلا بقسمة تكرارهم على المجموع الكلى لتلاميذ الفصل (اى $\frac{٣}{٦}$) أما النسبة المئوية فقد حصلنا عليها بضرب النسبة السابقة × ١٠٠ ، ولعلك لاحظت بعد هذا التحويل ان الذكور يمثلون نسبة أو (نسبة مئوية) أكبر من الاناث فى الفصل (ب) ، بينما تتساوى النسبتان فى الفصل (أ) وتدل بالطبع النسبة ٥ر أو النسبة المئوية ٥٠٪ على نقطة التوسط أو النزعة المركزية · وقد يستخدم فى قياس النزعة المنوال mode ويدل على أكثر القيم شيوعا فى التوزيع التكرارى ، وعلى الرغم من أنه مقياس مباشر للنزعة المركزية الا أنه أقل دقة من المتوسط والوسيط اللذين سنشير اليهما فما بعد ·

المدى المطلق :

المدى المطلق هو أحد مقاييس التشتت وأكثر بساطة وسهولة ، وهو فى جوهره مقياس للسعة ، ويحسب مباشرة بتحديد الفرق بين أكبر عدد وأقل عدد فى البيانات المتوافرة ، وهو بهذا المعنى يصلح للاستخدام مع المقاييس الاسمية ، فاذا كانت أرقام جلوس التلاميذ تمتد من ١ الى ٥٠٠ فان المدى المطلق فى هذه الحالة يصبح ٤٩٩ ، ويمكن استخدامه بنفس الطريقة مع درجات الاختبارات ·

وقد يستخدم فى قياس تشتت بيانات المقاييس الاسمية النسب والنسب

المئوية ايضا • فكلما ازدادت هذه النسب تطرفا بالزيادة أو النقص دل ذلك على درجة البعد عن نقطة التوسط (٥٠٪) •

معامل الارتباط بين بيانات المقاييس الاسمية :

يدل معامل الارتباط على مدى الاتفاق بين متغيرين ، أى ما اذا كانا يختلفان معها ، وفى أى اتجاه ، والى أى حد ؟ أو معنى اخر هو النزعة الى اقتران التغير فى ظاهرة بالتغير فى ظاهرة أخرى •

ويمتد معامل الارتباط بين (+ ١) والذى يعنى وجود ارتباط موجب كامل بين المتغيرين ، (فالصفر) الذى يعنى عدم وجـــود ارتباط بينهما واستقلالهما عن بعضهما تماما ، (- ١) الذى يعنى وجود ارتباط سالب كامل بين المتغيرين ، والواقع أن الحصول على معاملات الارتباط الكاملة ، سواء بالموجب أو السالب ، نادر الحدوث ان لم يكن مستحيلا • ولذلك فان أغلب معاملات الارتبـــاط بين المتغيرات فى المجالات النفســية والتربوية والاجتماعية عبارة عن كسر عشرى (أكبر من الصفر وأقل من الواحـــد الصحيح) يدل لى وجود علاقة جزئية ، وقد يكون هـــذا الكسر العشرى موجبا أو سالبا •

وتوجد طرق احصائية كثيرة لقياس العلاقة بين التكرارات المصنفة الى فئات المقاييس الاسمية ، وهى تنتمى الى ما يسمى الاحصاء الابارامترى ، ولا يتسع المقام لعرض هذه الطرق ويمكن الرجـــوع فى ذلك الى مرجع متخصص (٧) ، ونكتفى بعرض طريقة واحدة هى ما يسمى عامل ارتباط فاى •

مثال : حصل احد المعلمين على البيانات التالية من عدة فصول يقوم بالتدريس لها تتألف من الجنسين (الذكور والاناث) صنفوا جميعا فى ضوء نجاحهم أو فشلهم الدراسى كما يلى :

	التلاميذ	التلميذات	المجموع
ناجحون	(٣٥)	(٣٧)	٧٢
فاشلون	(١٤)	(٣٤)	٤٨
المجموع	٤٩	٧١	١٢٠

واراد هذا المعلم أن يحلل البيانات الاسمية فى الجدول السابق ليستخلص منه ما اذا كانت توجد علاقة بين الجنس والنجاح المدرسى . انه عليه فى هذه الحالة أن يحسب ماعمل ارتباط فاى ، وذلك بتحويل التكرار الثنائى للاعداد بين القوسين فى الجدول السابق الى اللغة الرمزية كما يلى :

التلاميذ	التلميذات
أ	ب
جـ	د

ثم تطبق المعادلة الاتية :

$$\text{ر ف} = \frac{\text{أ د} - \text{ب جـ}}{\sqrt{(\text{أ} + \text{ب})(\text{جـ} + \text{د})(\text{أ} + \text{جـ})(\text{ب} + \text{د})}}$$

وبالتعويض عن رمز المعادلة فان :

$$\text{ر} = \frac{(٣٥ \times ٣٤) - (٣٧ \times ١٤)}{\sqrt{(٧٢ \times ٤٨ \times ٥٩ \times ٧١)}} = +\ ١٩\text{ر}$$

وهو معامل ارتباط موجب ولكنه منخفض ، وقد لا تكـــون له دلالة احصائية (*) اى لا يدل على وجود علاقة بين الجنس والنجاح المدرسى ٠

تحليل بيانات مقاييس الرتبة

يمكن الحصول على اتجاه النزعة المركزية للبيانات الرتبية باستخدام الوسيط ، وعلى اتجاه التشتت باستخدام نصف المدى الربيعى ، وعلى معامل الارتباط كمقياس للعلاقة باستخدام معامل ارتباط الرتب ٠

الوسـيط :

يمكن تعريف الوسيط بأنه الرتبة التى تقسم التكرارات الى نصـــفين متساوين ، وفيما يلى مثال يوضح طريقة حسابه :

فئات الدرجات	الحدود الحقيقية للفئات	التكرار	التكرار المتجمع الصاعد
٥ ـ ٩	٤ر٥ ـ ٩ر٥	٤	٤
١٠ ـ ١٤	٩ر٥ ـ ١٤ر٥	١	٥
١٥ ـ ١٩	١٤ر٥ ـ ١٩ر٥	١	٦
٢٠ ـ ٢٤	١٩ر٥ ـ ٢٤ر٥	١٠	١٦ فئات الوسيط
٢٥ ـ ٢٩	٢٤ر٥ ـ ٢٩ر٥	٣	١٩
٣٠ ـ ٣٤	٢٩ر٥ ـ ٣٤ر٥	٥	٢٤
٣٥ ـ ٣٩	٣٤ر٥ ـ ٣٩ر٥	٣	٢٧
٤٠ ـ ٤٤	٣٩ر٥ ـ ٤٤ر٥	٠	٢٧
٤٥ ـ ٤٩	٤٤ر٥ ـ ٤٩ر٥	١	٢٨
	المجموع	٢٨	

(*) مفهوم الدلالة الاحصائية ينتمى الى الاحصاء الاستدلالى ، ولا يتسع المقام فى هذا الكتاب لشرحه ويمكن للقارىء الرجوع الى بعض المصادر المتخصصة فى الاحصاء النفسى والتربوى والاجتماعى ٠

ولحساب الوسيط من بيانات الجدول السابق يتبع الباحث الخطوات الآتية :

١ - تحديد رتبة الوسيط أى النقطة التى ينقسم عندها التكرار إلى نصفين بقسمة المجموع الكلى للتكرارات على ٢ أى أن رتبة الوسيط

$$= \frac{٢٨}{٢} = ١٤$$

٢ - دراسة التكرارات المتجمعة الصاعدة (العمود الأخير فى الجدول) لتحديد الفئة التى يقع فيها الوسيط ، وهى فى هذه الحالة الفئة (٢٠ - ٢٤)٠

٣ - تطبيق المعادلة التالية :

الوسيط = الحد الأدنى الحقيقى لفئة الوسيط +

$$\left[\frac{\text{رتبة الوسيط - التكرار السابق لفئة الوسيط}}{\text{تكرار فئة الوسيط}} \right] \times \text{مدى الفئة}$$

على النحو الآتى :

$$\text{الوسيط} = ١٩٫٥ + \frac{١٤ - ٦}{١٠} \times ٥ = ٢٣٫٥$$

نصف المدى الربيعى :

لعلك لاحظت أن الوسيط هو نقطة واحدة تقسم التكرار الكلى أى مجموع الأفراد أو الحالات الى نصفين ، ويمكن لهذه الفكرة أن تمتد منطقيا الى تقسيم التكرار الى أى عدد من الأقسام ، فمثلا يمكن أن ينقسم الى أربعة أقسام باستخدام ثلاث نقط التى تسمى فى هذه الحالة الربيعيات ، أو الى عشرة أقسام باستخدام تسع نقط تسمى العشيريات ، أو مائة قسم باستخدام ٩٩ نقطة تسمى المئينيات ، وتلتقى هنا بفكرة الربيعيات التى يعتمد عليها حساب نصف المدى الربيعى كمقياس احصائى لتشتت البيانات الرتبية٠

ولا يختلف حساب الربيع عن حساب الوسيط ، الا فى نقطة البدايا وعلى هذا يمكن حساب الربيع الأدنى أو الأول من بيانات الجدول السابق كما يلى :

١ ـ تحديد رتبة الربيع الأول بقسمة المجموع الكلى للتكرار على (٤) وضربها فى (١) أى أن :

$$\text{رتبة الربيع الأول او الأدنى} = \frac{٢٨ \times ١}{٤} = ٧$$

وهم ربع الأفراد الذين يفترض فيهم أن تقع درجاتهم أدنى من نقطة الربيع الأول أو الأدنى ، أما الأرباع الثلاثة الأخرى فتقع درجاتهم أعلى منها .

٢ ـ دراسة التكرارات المتجمعة الصاعدة لتحديد الفئة التى تقع فيها نقطة الربيع الأول وهى هذه الحالة الفئة (٢٠ ـ ٢٤) .

٣ ـ تطبيق المعادلة التالية التى لا تختلف فى صيغتها عن معادلة حساب الوسيط .

$$\text{الربيع الأول} = \text{الحد الأدنى الحقيقى لفئة الربيع الأول} + \left[\frac{\text{رتبة الربيع الأول} - \text{التكرار المتجمع السابق لفئته}}{\text{تكرار فئته}}\right] \times \text{مدى الفئة}$$

أى أن :

$$\text{الربيع الأول} = ١٩{,}٥ + \frac{٧ - ٦}{١٠} \times ٥ = ٢٠$$

أما الربيع الثانى فتتحدد من رتبته على النحو التالى بافتراض ن = المجموع الكلى للتكرار .

$$\text{رتبة الربيع الثانية او الأوسط} = \frac{ن \times ٢}{٤} = \frac{ن}{٢} = \text{الوسيط}$$

وهم نصف الأفراد الذين يفترض فيهم أن تقع درجاتهم أدنى من نقطة الربيع الثانى أو الوسيط ، وبالطبع فان النصف الثانى تقع درجاتهم أعلى منها .

ومعنى ذلك : ان الربيع الثانى هو الوسيط وسبق لنا حسابه ٠

اما الربيع الثالث أو الأعلى فيتحدد من بيانات الجدول السابق كمايلى:

$$١ ـ \text{رتبة الربيع الثالث أو الأعلى} = \frac{ن \times ٣}{٤} = \frac{٢٨ \times ٣}{٤} = ٢١$$

وهم ثلاثة أرباع الأفراد الذين يفترض فيهم أن تقع درجاتهم أدنى من نقطة الربيع الثالث أو الأعلى ، أما الربع الباقى فتقع درجاتهم أعلى منها

٢ ـ الفئة التى يقع فيها الربيع الثالث هى ٣٠ ـ ٣٤

$$٣ ـ \text{الربيع الثالث} = ٢٩{,}٥ + \left(\frac{٢١ - ١٩}{٥}\right) \times ٥ = ٣١{,}٥$$

ولحساب نصف المدى الربيعى كمقياس احصائى لتشتت البيـــانات الرتبية تعتمد على نصف الفرق بين الربيعيين الأول والثالث وهو ما يسمى نصف المدى الربيعى ٠

نصف المدى الربيعى =

$$\frac{\text{الربيع الثالث} - \text{الربيع الأول}}{٢}$$

$$= \frac{٣١{,}٥ - ٢٠}{٢} = ٥{,}٧٥$$

معامل ارتباط الرتب :

يستخدم فى القياس الاحصائى للعلاقة بين متغيرين من طبيعة رتبية

معامل ارتباط الرتب الذى ابتكره عالم النفس الانجليزى تشارلز سبيرمان ، ويوضح الجدول التالى طريقة حسابه ·

مثال :

حصل أحد الباحثين على ترتيب ١٥ تلميـــذا فى مـــادتى التاريخ والجغرافيا · حيث تدل الرتبة (١) على التلميذ الأول والرتبـــة (١٥) على التلميذ الأخير فى كل مادة على حدة ، مع ملاحظة أن التلاميذ الذين تساوت رتبتهم حصلوا على متوسط الرتب المتتالية ، أى ان التلاميذ د ، س حصلا فى التاريخ على رتبة ٥ر١٢ وهى متوسط الرتبتين ١٢ ، ١٣ ، والتلاميذ أ ، ز ، ن ، حصلوا جميعا على الرتبة (٨) وهى متوسط الرتب ٧ ، ٨ ، ٩ ·

التلاميذ	ترتيب التاريخ	ترتيب الجغرافيا	فروق الرتب (ف)	مربع فروق الرتب (ف٢)
أ	١١	٨	+ ٣	٠٠ر٩
ب	٤	٦	– ٢	٠٠ر٤
جـ	٩	٥	+ ٤	٠٠ر١٦
د	١٠	١٤	– ٤	٠٠ر١٦
هـ	٥ر١٤	١٥	– ٥ر٠	٢٥ر٠
و	٥ر١٤	١٢	+ ٥ر٢	٢٥ر٦
ز	٥ر١٢	٨	+ ٥ر٤	٢٥ر٢٠
ح	١	٣	– ٢	٠٠ر٤
ط	٣	١	+ ٢	٠٠ر٤
ى	٧	٤	+ ٣	٠٠ر٩
ك	٦	١٠	– ٤	٠٠ر١٦
ل	٢	٢	صفر	٠٠ر٠٠
م	٥	١٣	– ٨	٠٠ر٦٤
ن	٨	٨	صفر	٠٠ر٠٠
س	٥ر١٢	١١	+ ٥ر١	٢٥ر٢

مجـ ف = صفر مجـ ف٢ = ٠٠ر١٧١

ويحسب معامل ارتباط الرتب بالمعادلة الآتية :

$$ر ت = ١ - \left[\frac{٦ \times مجـ ف^٢}{ن (ن^٢ - ١)} \right]$$

$$= ١ - \frac{٦ \times ١٧١}{١٥ \times ٢٢٤} = ٠,٦٩٦$$

وهو معامل ارتباط موجب مرتفع نسبيا ، ويدل على وجود علاقة بين التحصيل فى كل من التاريخ والجغرافيا اى أن التلميذ الذى جاء ترتيبه مرتفعا فى التاريخ يحتمل أن يكون كذلك فى الجغرافيا والعكس صحيح ·

تحليل بيانات مقاييس المسافة والنسبة

يحصل الباحث على اتجاه النزعة المركزية للبيانات التى توفرها مقاييس المسافة والنسبة باستخدام المتوسط الحسابى ، وعلى اتجاه التشتت بحساب الانحراف المعيارى ، وعلى معامل الارتباط باستخدام طريقة بيرسون التى تسمى طريقة الارتباط التتابعى ، (أو طريقة حاصل ضرب العزوم) ·

المتوسط الحسابى :

المتوسط الحسابى كمقياس للنزعة المركزية لبيانات مقاييس المسافة والنسبة ، يعتمد على تقسيم الدرجات أو الكميات - وليس التكرارات كما هو الحال فى الوسيط - الى نصفين متساويين ، ويوضح المثال التالى ذلك :

مثال :

طبق أحد المعلمين اختبارا فى فهم القراءة على تلاميذ أحد فصول الصف الرابع الابتدائى فحصل على البيانات الآتية :

التلاميذ	أ	ب	جـ	د	هـ	و	ز	ح	ط
الدرجات	١	٣	٤	٥	٧	٥	٣	١	٧

وللحصول على المتوسط الحسابي كمقياس احصائي للنزعة المركزية يقوم المعلم بجمع درجات التلاميذ ، وهو في هذه الحالة ٣٦ ، وجمع عدد التلاميذ وهو في هذه الحالة ٩ ، ويحسب المتوسط الحسابي بالمعادلة الآتية :

$$\text{المتوسط الحسابي} = \frac{\text{مجموع الدرجات}}{\text{عدد الأفراد}} \text{ (أو) م} = \frac{\text{مج س}}{\text{ن}} = \frac{٣٦}{٩} = ٤$$

الانحراف المعياري :

يعتمد حساب الانحراف المعياري كمقياس احصائي لتشتت بيانات مقاييس المسافة والنسبة على حساب المتوسط الحسابي أولا ، ثم حساب الفروق أو الانحرافات عن هذا المتوسط ، ونظرا لأن مجموع هذه الفروق يساوي صفرا ، يستحيل علينا أن نحسب متوسط هذه الفروق ، ولذلك يلجأ الاحصائيون الى تربيع كل فرق أو (انحراف) للتخلص من العلامات الجبرية السالبة ، ثم نحسب بعد ذلك متوسط مربعات الانحرافات بقسمتها على عدد الأفراد ، ثم نستخرج الجذر التربيعي لمتوسط مربعات الانحرافات ، ويكون هذا الجذر التربيعي هو الانحراف المعياري .

ويمكن حساب الانحراف المعياري للمثال السابق ، وذلك بطرح المتوسط الحسابي (= ٤) ، من كل درجة من الدرجات السابقة فنحصل على الانحرافات عن المتوسط (ح) ، ثم نربعها (ح^{٢}) كما يلي :

التلاميذ	أ	ب	جـ	د	هـ	و	ز	ح	ط
ح	٣ـ	١ـ	صفر	١+	٣+	١+	١ـ	٣ـ	٣+
ح^{٢}	٩	١	صفر	١	٩	١	١	٩	٩

مجموع مربعات الانحرافات = ٤٠

ثم نحسب الانحراف المعيارى بالمعادلة التالية :

$$\text{الانحراف المعيارى} = \sqrt{\frac{\text{مجموع مربعات الانحرافات عن المتوسط}}{\text{عدد الأفراد}}}$$

$$\text{او ع} = \sqrt{\frac{\text{مج ح}^2}{\text{ن}}} = \sqrt{\frac{٤٠}{٩}} = ٢٫١١$$

معامل الارتباط التتابعى لبيرسون :

يستخدم فى حساب العلاقة بين مقاييس المسافة والنسبة الطريقة العامة فى حساب معامل الارتباط والتى ابتكرها عالم الاحصاء الانجليزى كارل بيرسون ، ويوضح المثال التالى طريقة الحصول على هذا المعامل ٠

مثال : طبق أحد المعلمين اختبارين يقيس احدهما تحصيل تلاميذ الصف الرابع من التعليم الأساسى فى مادة الحساب (س) ويقيس الاخر القدرة العددية (ص) لدى هؤلاء التلاميذ ، ويوضح الجدول الآتى البيانات التى حصل عليها وخطوات حساب معامل الارتباط ٠

التلاميذ	درجات اختبار التحصيل الحسابى (س)	درجات اختبار القدرة العددية (ص)	$س^2$	$ص^2$	س ص
أ	٢	٥	٤	٢٥	١٠
ب	٣	٧	٩	٤٩	٢١
جـ	٥	٦	٢٥	٣٦	٣٠
د	٧	١٠	٤٩	١٠٠	٧٠
هـ	٨	١٢	٦٤	١٤٤	٩٦
ن = ٥	مج س = ٢٥	مج ص = ٤٠	مج $س^2$ = ١٥١	مج $ص^2$ = ٥٣٤	مج س ص = ٢٢٧

ويمكن حساب معامل الارتباط مباشرة من بيانات الجدول السابق بتطبيق المعادلة الآتية :

$$\text{ر} = \frac{(\text{ن مج س ص}) - (\text{مج س} \times \text{مج ص})}{\sqrt{[(\text{ن مج س}^2) - (\text{مج س})^2] \times [(\text{ن مج ص}^2) - (\text{مج ص})^2]}}$$

وبالتعويض عن رموز المعادلة السابقة نجد أن :

$$\text{ر} = \frac{(٥ \times ٢٢٧) - (٢٥ \times ٤٠)}{\sqrt{[(٥ \times ١٥١) - ٦٢٥]\,[(٥ \times ٣٥٤) - ١٦٠٠]}} = \text{٩١ر}$$

وهو معامل ارتباط مرتفع جداً ، ويدل على ان التلميذ الذى يحصل على درجات عالية فى اختبار التحصيل الحسابى يحتمل احتمالا كبيرا انه يكون على درجة عالية فى اختبار القدرة العددية ، والعكس صحيح ٠

مراجع الباب الأول

١ ـ احمد زكى صالح : **الأسس النفسية للتعليم الثانوى** : القاهرة : دار النهضة العربية ، ١٩٧٢ .

٢ ـ احمد زكى صالح : **علم النفس التربوى** : القاهرة : النهضة المصرية ، ١٩٧٢ .

٣ ـ جابر عبد الحميد : **التقويم التربوى والقياس النفسى** . دار النهضة العربية ، ١٩٨٣ .

٤ ـ سعيد يامشمرس ، السيد محمد خيرى ، يحيى مهنى : **التقوية التربوى** الرياض : منشورات دار الفيصل الثقافية ، ١٩٨٠ .

٥ ـ صلاح الدين محمود علام : تحليل البيانات فى البحوث النفسية والتربوية . دار الفكر العربى ، ١٩٨٥ .

٦ ـ رايتسون ، ج . جاستمان ، ج . و . روبنز ، (ترجمة : محمد محمد عاشور وآخرين) : **التقويم فى التربية الحديثة** . القاهرة : الانجلو المصرية ، ١٩٦٥ .

٧ ـ رمزية الغريب : القياس اللابرمترى . مكتبة الأنجلو المصرية ، ١٩٨٥ .

٨ ـ رمزية الغريب : **التقويم والقياس النفسى والتربوى** . القاهرة : الأنجلو ١٩٧٠ .

٩ ـ عبد الله السيد عبد الجواد : المؤشرات التربوية واستخدام الرياضيات فى العلوم الانسانية ، أسيوط ، ١٩٨٣ .

١٠ ـ فؤاد أبو حطب . **القدرات العقلية** . القاهرة : الأنجلو المصرية ط ٤ ١٩٨٣ .

١١ ـ فؤاد أبو حطب ، آمال صادق ، علم النفس التربوى ، الأنجلو المصرية، (ط ٣) ، ١٩٨٤ .

١٢ ـ فؤاد أبو حطب ، بديوى علام : الاحصاء النفسى والتربوى ، مكتبة الأنجلو المصرية (تحت الطبع) .

١٣ ـ فؤاد البهى السيد : علم النفس الاحصائى وقياس العقل البشرى . دار الفكر العربى ، ١٩٧٨ .

١٣ أ ـ محمد خليفة التونسى : التقويم والتقييم ، مجلة العربى ، يونيه ١٩٨٦

١٣ ب ـ محمد خليفة التونسى أضواء على لغتنا السمحة، كتاب العربى، ٨٦

١٣ جـ ـ محمد خليفة التونسى : الاشتقاق من المشتق ، مجلة العربى، يولية ٨٦

١٤ ـ محمد عبد السلام احمد : **القياس النفسى والتربوى** . القاهرة النهضة المصرية ، ١٩٦٠ .

١٥ ـ نعيم عطية : التقييم التربوى الهادف ، بيروت : منشورات دار الكتاب اللبنانى ، ١٩٧٠ .

١٦ ـ يوسف محمود الشيخ · جابر عبد الحميد جابر : **سيكولوجية الفروق الفردية** · القاهرة : النهضة العربية ١٩٦٤ ·

17. Allport, G.W. **Pattern and growth in personality,** New York: Holt Rinehart, 1961, (1937).
18. Anastasi, A. Psychological testing New York, Macmillan Publishing, (5th ed.) 1982.
19. Anastasi, A. **Differential psychology** (3rd ed.), New York : Macmillan, 1953.
20. Bloom, B. **Taxonomy of educational objectives.** London : Longmans, 1956.
21. Bordon, E.S. **Psychological counseling.** New York : Appleton, 1955.
22. Cook, W.W. The Function of measurement in the faciliitation of Learning. In Linquist, E.F. (ed.) **Educational Measurement** Amer. Council on Educ., Washington, D.C., 1951, 3-64.
23. Coombs, C.H. Theory and méthod of social measurement. In : Festinger, L. and Katz (eds.) Research methods of behavioral Sciences. New Delhi Ed., 1970.
24. Cronbach, J. L. **Essentials of psychological testing,** New York : Harper (4th ed.), 1984.
25. Gronlund, **Measurement and evalution in teaching.** New York : Macmillan, (4th ed.), 1981.
26. Krathwohi, D.R., et al. **Taxonomy of educational objectives.** II Affective **domeuin,** 1964.
27. Nunnally, J.C. (Jr.) Introduction to psychological measurement. New York : McGraw-Hill, 1970.
28. Rogers, C.R. **Counseling and psychotherapy.** Cambridge, Mass. : Riverside Press, 1942.
29. Rogers, C.R. **Client-centered therapy. Cambridge,** Mass.: Riverside Press, 1951.
30. Stevens, S.S. Mathematics, measurement and psychophysics. In : Stevens, S.S. (ed.) **Handbook of Experimental Psychology.** New York : Wiley, 1951.
31. Thorndike, R. and Hageu, E. **Measurement and evaluation in psychology and education.** New York : Wily, 1955.
32. Tylers, R.W. The Function of measurement in improving instruction. In: Lindquist, E.F. (ed.) **Educational Measurement,** 1951, 47-67.
33. Tyler, L.C. **The work of the counselor** (2nd ed.). New York York : Appleton-century, 1961.

البـــاب الثانى

شروط الاختبارات النفســية

الفصــل الثالث

فن اعطاء الاختبارات

ماذا يقصد بتقنين الاختبار ؟

أشرنا في الفصل الأول الى تعريف للقياس النفسي يشير الى ضرورة التعبير الصريح عن قواعده، وهذا ما يتضمنه مفهوم «التقنين Standardization». فحين نقول ان هذا القياس « مقنن » فان ذلك يعنى فى جوهره أنه لو استخدمه أفراد مختلفون يحصلون على نتائج متماثلة · ويتطلب هذا بالطبع توحيد اجراءات تطبيق الاختبار وتصحيحه ، فاذا كانت الدرجات التى يحصل عليها الأفراد المختلفون قابلة للمقارنة فان شروط الاختبار وظروف اعطائه يجب أن تكون موحدة للجميع · وهذا الشرط ما هو الا تطبيق خاص للحاجة الى الشروط والظروف المضبوطة فى جميع الملاحظات العلمية · ومثل هذا التقنين يجب أن يشمل المواد المستخدمة وحدود الزمن والتعليمات الشفوية والتحريرية التى تعطى للمفحوصين والأمثلة التوضيحية ، وطرق تناول أسئلة المفحوصين واستفساراتهم ، وكل ظرف أو حالة لا ترتبط بما يتناوله الاختبار ويمكن أن تؤثر فى أداء المفحوصين فى الاختبار كلما كان ذلك ممكنا ·

ويوفر تقنين قواعد القيـــاس للباحثين فى علم النفس ظـرفا ملائما لسهولة «الاتصال» ، لأن من أهم خصائص العلم وجود نوع من الاتصـــال الجيد بين العلماء بحيث يستطيع الباحث أن يقارن بين نتائجه ونتائج زملائه الذين يتناولون نفس المشكلات بالدراسة · فاذا افترضنا أن أحــــد الباحثين أجرى تجربة على آثار التوتر فى استجابات القلق وسجل نتائجه فى عبارة مثل أن « التوتر يؤدى الى ظهور القلق عند الأطفال » فان مثل هذه النتيجة لا تفيد فى « الاتصال » الجيد لأن الباحثين الآخرين قد يختلفون حـــول ما يقصده « بظهور القلق » ، بل قد يصعب عليهم ان لم يستحل ، تكرار نفس التجربة للتحقق من صحة الفرض ، ويمكن تحقيق نوع من الاتصال الأكفأ لو استطاع الباحث أن يلاحظ القلق باختبار « مقنن » أى محددة قواعد اعطائه وشروط تطبيقه ·

وتتأثر عملية اعطاء الاختبار بعوامل متعددة ، منها ما يتصل بمن يعطى الاختبار وما يفعله ، ومنها ما يتصل بمن يقوم باداء معين على هذا الاختبار وما يحيط به من ظروف ومؤثرات ، ومنها ما يتصل بالاختبار نفسه وما يتميز به من خصائص وما يقصد الى أن يقيس · واذا كانت هناك عوامل وشروط ينبغى أن تتوفر فى الاختبار نفسه حتى يكون اداة قياس يعتمد عليها ويطمأن اليها ، وهى ما سنتناوله فى الفصول التالية ، فان هناك شروطا وظروفا تؤلف ما نسميه فن اعطاء أو تطبيق الاختبارات على الفاحص أن يراعيها ويحققها، وهى التى من شأنها أن توفر ، أول ما توفر ، الجــو النفسى ــ الاجتماعى المناسب للمفحوص بحيث يأتى اداؤه فى الاختبار ، كما اراد واضعه ، وكما يريد الفاحص ، دقيقا وصادقا ، ليكون أساسا متينا لقرار تزيد احتمالات دقته وجدواه ·

ولعل مسألة تقنين الاختبارات وغيرها من أدوات التقويم النفسى هى التى جعلت علماء النفس على درجة كبيرة من الحساسية لمســألة اخلاقيات المهنة · وقد ظهر هذا الاهتمام لأول مرة فى أوائل الخمسينات حين أصدرت الجمعية الأمريكية لعلم النفس مجموعة من المعايير الأخــــلاقية التى يجب على الباحثين السيكولوجين مراعاتها (١٢) ثم سرعان ما اتبعت ذلك باصدار كتيب التوصيات الفنية الخاصـــــة بالاختبارات النفسية والأســـــاليب التشخيصية (١٣) ، وقد تطورت هذه الوثائق وتعدلت طوال السنوات الثلاثين الماضية وصدرت من كل منها طبعات جديدة (١٤ ، ١٥ ، ١٦ ، ١٧) · ونحن فى حاجة الى ترشيد هذه الخبرة والممارسة السيكولوجيتين فى مصر ، وقد عرض هذه الضرورة محمد خليفة بركات (١٠) فى المؤتمر الثانى لعلم النفس فى مصر عام ١٩٨٦ ، ولعل الجمعية المصرية للدراسات النفسية تقوم بهذه المســـئولية ·

واجبات الفاحص ومسئولياته فى اعطاء الاختبار :

١ ــ **التدريب على اعطاء الاختبار** : من الاختبارات ما يسهل اعطاؤه فلا يحتاج الى تدريب خاص للفاحص ، ومنها ما يحتاج الى تدريب ومران قد يمتد بضعة شهور · والاختبارات الفردية بصفة عامة تحتاج الى تدريب أكثر من الاختبارات الجماعية ، وان كانت هناك بعض الاختبارات الجماعية التى تتطلب عناية كبيرة فى اعطائها · وكذلك فان الاختبارات الاسقاطية تحتاج

الى تدريب خاص قد يصل فى بعض الجامعات الى الحصـــول على دبلومة متخصصة فيها ، وهو ما لا تحتاجه الاستخبارات او الاستبيانات ، وتحتاج المقابلة أو الملاحظة تدريبا متخصصا لا تحتاجه بعض وسائل التقرير الذاتى٠ واذا كانت كل مسئولية الفاحص تنحصر فى قراءة مجموعة من التوجيهات والتعليمات المطبوعة ، فان اى شخص ذا ضمير يقظ حســـاس ، وليس فى شخصيته ما يحدث تهديدا للاخرين ، يمكن ان يقوم بها بنجاح ٠ اما اذا كانت الضرورة تقتضى ان يسأل المفحوص فرديا ، وان تستعمل أسئلة تتبعية اذا كان السؤال الأول غير واضح ٠ فان هذا يقتضى توفر قدر كبير من الخبرة والمهارة ، اى تدريبا خاصا وطويلا بما يتفق مع درجة صـــعوبة الموقف الاختبارى ، ودقة القرار الذى يتخذ فى ضوء بيانات الاختبار ٠

وتشير بعض المؤسسات المسئولة عن طبع الاختبارات وتوزيعها ، وعلى الخصوص NFER فى انجلترا ETS(*) فى أمريكا أن تتوافر فيمن يشترى الاختبارات شروط ومؤهلات معينة ـ قد تصل الى الدكتوراه فى علم النفس قبل السماح ببيعه أو توزيعه ، ناهيك عن القيود الصارمة على التوزيع والنشر التى تفرضها هذه المؤسسات وغيرها على الاختبارات حتى لا تفقد وظائفها الأساسية فى التقويم النفسى ٠ ولك أن تقارن هـذا بما يشـيع فى بلادنا من فوضى حيث تباع الاختبارات وتشترى ، بل وتصور ، دون رقيب ، بل يصل الأمر الى حد اعادة طبعها كاملة فى ملاحق رســـائل الماجستير والدكتوراه وفى هذا كله خروج على الميثاق الأخلاقى السيكولوجى لاستخدام الاختبارات ٠

٢ ـ **الفة الفاحص بالاختبار** : من اهم الواجبات التى ينبغى ان يراعيها الفاحص أن يتيح الفرصة لكل مفحوص ليظهر قدرته ، وليحصل على نتائج يمكن مقارنتها بنتائج مفحوصين آخرين ، فعندما تقوم عملية انتقاء ما على نتائج اختبار من الاختبارات فانها لن تكون عملية عادلة مالم يتم اختبـــار الأشخاص جميعا بطريقة موحدة ومتشابهة تماما ٠

ولهذا كان من الضرورى أن يعرف الفاحص الاختبار الذى سيستعمله، بل ويألفه الى أبعد درجة ممكنة ، اذا كنا نهدف الى تحقيق عملية اختبـــار

(*) همـا
(NFER) National Foundation for Educational Research
(ETS) Educational Testing services _ على التوالى ٠

بقيقة منصفة . ومما يساعده على هذه المعرفة والألفة أن يقرأ كراسة التعليمات التى تصحب الاختبار عادة بعناية وان يلتزم بما فيها التزاما تاما .

٣ - **الالتزام بالاتجاه العلمى** : على الفاحص ان يلتزم اتجاها علميا محايدا اثناء عملية الاختبار ، ذلك لأن المختبر غالبا ما يكون مهتما بالمفحوصين ، وقد يميل الى أن يراهم يحسنون صنعا فى الاختبار الذى يعطيهم اياه ، فينزع الى أن يساعدهم بنية حسنة ، بطريق مباشر او بطرق خفية ، وهذا أكثر حدوثا فى مواقف الاختبار الفردية . ولكن على المختبر أن يلاحظ هذا فى نفسه ، وأن يتجنبه . ولا شك ان مما يساعد على هذا ان يلتزم المختبر بالتعليمات المصاحبة للاختبار ، كما اشرنا ، فاذا ذكرت التعليمات مثلا أن سؤالا ما يكرر مرة واحدة . فلا يكرره المختبر اكثر من مرة واحدة فى أى ظرف من الظروف .

وكذلك على المختبر أن يلاحظ اى تعبيرات للوجه او كلمات التشجيع او الاشارات التى يراد بها حث همة المفحوص ، هذا كله يضعف من قيمة الاختبار ومن تقنينه . وعلى هذا فمن المهم بالنسبة للفاحص ان يجعل هيئته وسمته وتعبيرات وجهه واشاراته غير كاشفة . على ألا يعطى هذا المفحوص احساسا بأن المختبر لا يكترث به ، بل لابد ان يحرص المختبر على ان يوحى للمفحوص ويؤكد له فى صمت اهتمامه بكل ما يقول .

٤ - **المحافظة على العلاقة بين الفاحص والمفحوص** : يلفتنا هذا الى اهمية المحافظة على العلاقة بين الفاحص والمفحوص فى موقف الاختبار . فاذا كان مطلوبا من المفحوص أن يظهر افضل ما عنده فى الاختبار ، اى انه ينبغى ان يظهر باستمرار رغبته فى التعاون مع المختبر فانه من الضرورى أن تقوم علاقة انسانية تتسم بالاطمئنان والثقة والقبول المتبادل بين طرفى هذه العملية - الاجتماعية فى أساسها - أى علاقة المفحوص والمختبر .

ونسارع فنشير الى أنه ليست هناك قواعد محددة ولا نصائح تتبع لاقامة ذلك النوع من العلاقة بين الفاحص والمفحوص ، ولكنا نستطيع ان نقول بصفة عامة ان الشخص الذى يألف الناس ويألفونه ، والذى يوحى بالثقة وبالطمأنينة فى نفوس الآخرين ، يستطيع ان يجد من الوسائل ما يساعده على انشاء تلك العلاقة بسماتها التى حددناها .

وهناك علامات تدل على سوء العلاقة الاختبارية ، أو على أن الفاحص غير موفق فى اقامة العلاقة المطلوبة بمن يختبرهم ، ونذكر بعضا من هذه العلامات حتى يستطيع الفاحص الفطن أن يتعرف عليها ويدركها ثم يستجيب لها بما يؤدى الى اختفائها أو الاقلال منها · من هذه العلامات عدم انتباه المفحوصين أثناء القاء التعليمات ، أو الانتهاء قبل أن يحين الموعد المحدد ، أو عدم الاستقرار ، أو التماس أخطاء فى الاختبار · أو اظهار اليأس من امكان الأداء أو العمل فى الاختبار · ·

٥ - مسئولية الفاحص الأولى اختيار الاختبار الملائم للمفحوص : من الممارسات السيئة والخطيرة أن يعطى الفاحص للمفحوص اختبارا للقدرة مثلا لا يناسبه ثم يؤدى فيه أداء سيئا لأسباب غير مرتبطة بالأداء الاختبارى ذاته · فاختبار المتاهة مثلا لا يجب ان يعطى لمفحوصين يعانون من نقص فى التحكم فى حركات اليد ، وبالمثل لا يجب أن يعطى اختبار للميول المهنية يلائم طلاب الجامعات لطلاب مرحلة التعليم الأساسى ما لم يكن قد اجرى تقنين جديد له يؤكد صلاحيته للمستوى التعليمى الجديد ، ويشـــمل ذلك بالطبع التأكد من أن تلاميذ مرحلة التعليم الأساسى يفهمون تعليماته وأسئلته ·

ومن المعلوم فى علم القياس النفسى وفنونه التطبيقية أن حصـــول المفحوص على درجة مضللة فى اختبار غير ملائم اسوأ من عدم الحصول على هذه الدرجة على الاطلاق · ومسئولية اختيار الاختبار الملائم للمفحوص تظل فى النهاية مسئولية الفاحص · وتبقى مسئوليته هذه حتى ولو طلب منه تطبيق اختبار بذاته من سلطة خارجية فعليه أن يستخدم مهارته وكفاءته فى الحكم على ملاءمة الاختبار، ويصدق هذا على مساعدى الباحثين الذين يقومون بجمع البيانات · ولعلنا فى هذا الصدد لا نبالغ اذا قلنا ان بحثا أو دراسة أو فحصا يعتمد على بيانات زائفة أو سيئة أو مضللة اسوأ للعلم وأضـــل للممارسة العملية من عدم اجراء بحوث على الاطلاق ·

طريقة اعطاء الاختبار :

١ - ظروف عملية الاختبار : هناك ظروف مادية لابد أن تتوفر فى الموقف الاختبارى لما لها من آثار عميقة على سلوك الأشخاص من الفاحصين وبالتالى

على نتائج اختباراتهم ، من هذه الظروف التهوية والاضاءة ولهما أثر واضح على أداء الأشخاص فى الاختبارات ، سواء كان هذا الأداء كتابة أو عملا . وكذلك من الضرورى توفير أماكن جلوس تساعد على القيام بهذا الأداء ، من حيث الاتساع وامكان وضع الأدوات على مناضد بحيث يسهل تناولها واستعمالها وهكذا .

أما ترتيب أماكن الجلوس فلابد أن يراعى فيه أن يساعد المفحوصين على الاستماع الجيد والسليم للارشادات والتوجيهات . وهنا نشير الى أن الصالات الكبيرة الواسعة لا تكون مناسبة للاختبارات الجماعية الا اذا توفر العدد الكافى من المشرفين على أن يوزعوا توزيعا مناسبا . وحتى مع توفر العدد الكافى من المشرفين فلا تزال هناك صعوبة لابد أن نتغلب عليها وهى أن المفحوصين فى صالة ضخمة قد يحجمون عن السؤال أو الاستفسار من المشرفين خوفا ورهبا ، وعلى هذا ، ينبغى عند اجراء اختبار جماعى فى صالة كبيرة أن يشجع المفحوصون على أن يسألوا ويستفسروا عما يعن لهم أو يلتبس عليهم فيما يتصل بما بين أيديهم من اختبار .

أما اذا انتقلنا الى ظروف المفحوص نفسه ، من الناحية الجسمية والنفسية ، فيجب أن نعطيها قدرا كبيرا من الاهتمام . وأول ما يلاحظ هو اختيار الوقت المناسب له من الناحية الجسمية والنفسية ، لأن اعطاء الاختبار فى ظروف جسمية أو انفعالية غير مناسبة ، لا تجعل اداءه عينة صادقة وممثلة لسلوكه .

ومما يساعد على تحقيق هذه الظروف توفير قسط مناسب من الراحة قبيل اعطاء الاختبارات ، ثم أثناءها بصورة منظمة دون ارتباك أو فوضى . وكذلك مراعاة أن لا تعطى الاختبارات عقب فترة من المجهود الجسمى أو العقلى الشاق أو عقب التعرض لحالة انفعالية حادة ، الا اذا كان الغرض من القياس ذلك .

أما اذا اضطررنا ، لسبب ما ، أن نعطى اختبارات فى ظروف غير مناسبة فيجب أن نعتبر نتائجها أولية وأن ننظر اليها نظرة ناقدة وأن نأخذها بغير قليل من التحفظ والحذر .

٢ ـ ادارة الجماعة وضبطها : لا يعطى الاختبار الا لأفراد نتوقع أن

يفهموا توجيهات وتعليمات معينة ، كما نتوقع أن يسلكوا وفق هذه التوجيهات والتعليمات التى يوجهها الفاحص اليهم ، ولذا كان أساس الاختبار الجماعى، من ناحية الفاحص ، هو ضبط الجماعة وادارتها بطريقة تمكنه شخصيا من اداء واجباته المتصلة باعطاء الاختبار بدقة وحكمة ، ثم تمكن الجماعة من اظهار أحسن اداء ممكن على هذا الاختبار ٠

ومما يحقق للفاحص النجاح فى هذا أن يقف أمام الجماعة بحيث يرى كل أعضائها وبحيث يستطيع كل أعضائها أن يروه ويسمعوه ، ويلقى تعليماته بعد أن يستقر ويسكن الجميع فى أماكنها ، ويلاحظ أن تكون التعليمات بسيطة واضحة وأن يلقيها الواحدة اثر الأخرى بهدوء وأناة بحيث يكون كل جزء وكل فقرة من التعليمات متميزة ليتجنب سوء الفهم والخلط من جانب المفحوصين ٠ واذ كان انتباه الجماعة له كاملا ، وهو ما ينبغى أن يحرص عليه ، فانه لن يحتاج الى تكرار شىء مما يقول ٠

ومن الممكن ان يلجأ الفاحص الى التمثيل ليزيد التعليمات وضوحا ، مالم يكن هناك نص فى دليل الاختبار او تعليمات موجهة الى الفاحص نفسه تنص على منع التمثيل ، ففى هذه الحالة عليه أن يتجنب التمثيل والتوضيح ٠ ومن امثلة التوضيح مثلا انه اذا كانت التعليمات تشير الى قلب الصفحة الأولى من الاختبار ، فلا بأس من أن يمسك المختبر الاختبار عاليا ثم يقلب الصفحة ، او ان يشير الى المكان الذى ينبغى أن تكتب فيه البيانات الخاصة بالمفحوص ، وهكذا ٠

ومن المهم أن يكون واضحا أن للمفحوصين الحق دائما فى ان يستفسروا ويسألوا ، ولعله من الأفضل ان يتوقع المختبر الأسئلة ويحاول الاجابة عنها قبل أن تسأل ٠

هل يتحقق ضبط الجماعة المفحوصة ويتيسر اجراء الاختبار فيها عن طريق اتباع اسلوب صارم جامد من الفاحص ؟ لا شك ان مثل هذا الأسلوب يجعل امر ضبط الجماعة وادارتها سهلا لا مشقة منه ولا عناء على الفاحص الا أن مثل هذا الأسلوب يخلع على جو الاختبار طابعا غير انسانى ، ومن المهم أن لا ننسى فى الموقف ا لاختبارى أو فى أى موقف اجتماعى بل وانسانى

بصفة عامة ، أنه اذا كان ضبط الجماعة ضرورة لتسيير امـــورها وتحقيق مطالبها ، فان من الضرورى أيضا أن يحس اعضاء الجماعة بأن من يوجههم أو يشرف عليهم انسان يهتم بهم ويحرص عليهم وليس شخصا رسميا يقيم بينهم وبين نفسه أستارا تجعلهم ينفرون منه مما يمتد أثره الى ما بين أيديهم من عمل يطلب منهم أداؤه · والموقف الاختبارى على أى حال ، لا يتطلب الا قدرا من الضبط يحققه قدر يسير من الرســـمية فى العـلاقات بين المفحوصين والفاحص · وقد تزيد درجة الرسمية هذه أو تنقص حسب أغراض الاختبار وظروفه، ففي بعض الاختبارات التحصيلية ـ وخاصة في مواقف الامتحانات ـ قد يحتاج الأمر لدرجة كبيرة من الرسمية والا حدثت الفوضى وشاع الغش ، ولعل ما يحدث فى وقتنا الحاضر فى الامتحانات المدرسية من مشكلات نظـام يؤيد هذا القول ، وهو على اية حال أمر يحتاج لدراسة مستقلة ·

ومن المهم أن نشير الى أن الفاحص لابد أن يرجع الى فطنته وتقديره ويعتمد على حصافته فى كثير من الأحيان أثناء الموقف الاختبـــارى ، فاذا نشأت ظروف اضطرارية أثناء الاختبار ، كأن يمرض مفحوص فجأة ، مثلا ، فهنا لابد أن يحتكم الفاحص الى تقديره الخاص ، مراعيا ان الهدف من أى اختبارهو الحصول على معلومات مفيدة ودقيقة عن الشخص وله ، وليس مجرد اجراء اختبار كأنه عمل روتيني والفراغ منه على أى حال وبأية صورة ·

٣ ـ **التعليمات للمفحوص** : ان اعطاء التعليمات والارشـــادات أهم مسئولية يتحملها الفاحص ، ذلك لأنه من الضرورى فى الاختبارات المقنة أن يعطى المختبر التعليمات كما وردت فى الاختبار دون زيادة أو نقص ، فعلى الرغم من أن الاختبارات تقيس الفروق الفردية فان طرقها ووســائلها تكون مرسومة بحيث تلغى التفاوت والاختلاف والتباين فى كل خاصية من خواص المفحوص فيما عدا تلك الخاصية التى يريد أن يقيسها الاختبار بالفعل · وعلى هذا فانه من الأهمية بمكان بالنسبة للفاحص أن تتساوى جميع الظروف ـ قدر الامكان ـ بين المفحوصين فيما عدا ذلك العامل أو تلك الخاصية التى يريد الاختبار قياسها ·

وعادة ما نسمح التعليمات نفسها بالأسئلة والاستفسارات ، ولكن عندما يجيب الفاحص على الأسئلة عليه الا يتعرض لتوضيح شيء أكثر مما فى

التعليمات ، اذ عليه أن يتذكر دائما أن نتائج مفحوصية قد تقـــارن بنتائج مفحوصين آخرين فى أماكن وأوقات أخرى ٠

وعلى أى حال فان الاختبار الجيد يأخذ فى الاعتبار كما أشرنا أكبر قدر من الأسئلة المتوقعة ويرتب اجابات وحلول لها حتى حتى تكون طريقة اعطائه مقننة، وبدون ذلك فان هذه الأسئلة قد يفهمها البعض على وجه بينما يفهمها البعض الآخر على وجه آخر ، ومن الضرورى دائما تحديد المطلوب من المفحوصين بصورة واضحة محددة الى أبعد حد مستطاع ٠

وفى بعض الأحيان تكون طريقة اتباع المفحوص التعليمـــات المقننة مقصودا بها أن تؤثر فى درجته ، فاذا أعطيت هذه التعلميات بطرق مختلفة ترتب على ذلك اختلافا فى الأداء لم يكن فى الحسبان ٠

ومن المسائل الهامة التى تتصل بتقنين تعليمات الاختبار ما يتصـــل بالتخمين ٠ ويظهر التخمين بصورة أقوى وأوضح فى اختبارات القدرة واختبارات التحصيل حيث يسأل المفحوص : « اذا لم أعرف الاجابة فهل من الممكن أن ألجأ الى التخمين ؟ » وفى بعض الاحيان تتضمن توجيهـــات الاختبارات نصوصا تبين الاجابة على مثل هذا السؤال ، ولكن حتى مع وجود مثل هذه النصوص فان الأمر لا يخلو من ابهام يتطلب الوضوح ٠ وانه مما يتعارض مع أصول اعطاء الاختبارات أن يتطوع الفاحص باعطاء نصح وتوصية من عنده لم تنص عليه توجيهات الاختبار ٠

ما هو الموقف السليم بالنسبة للتخمين فى الاختبارات ؟ يمكن أن نلخص ما ينبغى أن نعمله فى مواجهة التخمين فيما يلى ٠

(أ) على الفاحص أن يؤكد ما تشير اليه تعليمات الاختبار ، فاذا كانت هذه التعليمات تشجع التخمين أكده وشجع عليه أما اذا كانت تنص على عدم الاعتماد على التخمين فلعليه أن يظهر هذا ٠

(ب) أما اذا لم تكن فى دليل الاختبار اشـــارة الى التخمين فعلى الفاحص أن ينصح المفحوصين بأن يستعملوا ذكاءهم وتقديرهم فيما يختص بما اذا كان من الممكن أن يخمنوا ، عندما تواجههم أســـئلة أو عناصر من الاختبار لا يكونون واثقين من صحة الاجابة عنها ، بحيث تكون تخميناتهم

مستنيرة ، فلا يختارون الا ما يعتقدون أنه أقرب الى الصواب ، ولا تكون المسألة كلها فى يد الصدفة والتخمين غير المعقول ·

(هـ) من الضرورى أن يتعلم أولئك الذين سوف تطبق عليهم الاختبارات وخاصة من الطلاب ، أن الاسراف فى التخمين فى اختبـــــارات القدرة واختبارات التحصيل ليس فى صالحهم على الاطلاق ·

(د) من المعروف أنه قد وضعت معادلات لتصحيح أثر التخمين فى الاختبارات ، وعلى الفاحص أن يلجأ الى استخدام المعادلة المناسبة عنـــد تصحيح الاختبارات ليفسر الدرجة التى يحصل عليهـــا المفحوص تفسيرا معقولا ·

استثارة الدوافع لأخذ الاختبار :

هناك حقيقة هامة على كل مشتغل بالقياس النفسى أن يعيها وأن يعمل وفقا لها أولا وهى أن الشخص الذى نختبره لابد أن يهتم بنتائج الاختبار الذى يعطى له وأن تكون لهذه النتائج قيمة عنده · ففى القياس النفسى – الذى يختلف عن القياس المادى – نجد أن المفحوص يضع نفسه فى المقياس ، ومالم يهتم شخصيا بهذه العملية فاننا لا نستطيع قياسه وان لم يتحقق هذا فاننا لا يمكن اختباره بدرجة نطمئن فيها على ما نحصل عليه من بيانات عنه · واذا كانت الدوافع لا تمثل جانبا له أهمية فى أنواع القياس المادى المختلفة ، مثل الطول والوزن ، فانه لها أهميتها الكبرى فى القياس النفسى ، وحتى يكون القياس النفسى دقيقا لابد أن يقبل المفحوص على الاختبار ويتحرك نحوه · ولهذا كانت استثارة دوافع المفحوص للاختبـــار مسألة جديرة باهتمـــام الفاحصين وعنايتهم ·

(١) **الحوافز التى ترفع الدرجة** : الحوافز التى تساعد على الأداء فى الاختبارات النفسية متعددة ومتنوعة ، فمنها مثلا الرغبة فيها ، ومنها ذلك الحماس الذى ينتج من اندماج الذات والاستغراق الذى يعبر عن الرغبة فى الاحتفاظ باحترام الذات واحترام الآخرين · ومن هذه الحوافز أيضا مجرد حب العمل والاشتغال بحل عناصر اختبار ما ، فمثل هذا الحب والاهتمام

يستثير الشخص الى بذل مجهود أكبر فيه · ولا نستطيع أن نغفل حافزا يقوم بدور هام فى بعض الأحيان فى دفع المفحوصين وتحريكهم للاداء على اختبار ما ، ذلك هو الرغبة فى اطاعة السلطة وتنفيذ التزامها كما تتمثل فى الفاحص وما يوجهه الى المفحوصين من تعليمات وتوجيهات ، وهذه الرغبة كما هو واضح متصلة بحاجات اجتماعية عميقة منها مثلا الحاجة الى الاعتراف والتقدير ، وكذلك الحاجة الى الضبط والنظام ·

وتظهر الحاجة الى استثارة حوافز المفحوص ورفع دوافعه الى الدرجة الملائمة فى اختبارات القدرة واختبارات التحصيل بصفة خاصة ، وذلك لأن أهم ما يعنينا من هذين النوعين من الاختبارات هو الحصول على احسن اداء ممكن عليه من المفحوصين ، أو الحصول على أعلى انتاجية ممكنة فيهم اذا استعملنا تعبيرا ذا مسحة اقتصادية ·

والذى لا شك فيه أن مراعاة ما ذكرناه ، وما سناقش فيما يلى من شروط وظروف فى الموقف الاختبارى ، مما يساعد على تحقيق ذلك المستوى الأنسب من الدوافع لأخذ الاختبار والاستمرار فيه بجهد متوازن مناسب· وقد بذلت محاولات متعددة لاستعمال انواع بسيطة من المكافآت لرفع مستوى الأداء على اختبارات القدرة بنوع خاص ، ومن امثلة هذه المكافآت البسيطة الجوائز ، والنقود ، وأحاديث التشجيع ، وقد أشارت الأبحاث التى أجريت فى هذا الصدد (١٨ ، ٢٢ ، ٢٥) أن مثل هذه المكافآت لم تحدث أى تحسن فى درجات اختبارات القدرة عن الدرجات التى حصل عليها المفحوصون الذين اخذوا الاختبارات نفسها فى الظروف العادية لاعطاء الاختبارات · ومما يلاحظ انه من الممكن أن تتحسن الدرجات على الاختبار عندما تحول دوافع الشخص من التركيز على مثل تلك المكافآت البسيطة الى الاهتمام بدرجات الاختبار التى يحصل عليها ·

وحتى فى الاختبارات التى تعطى لأغراض البحث العلمى وليس لاتخاذ قرار بشأن المفحوصين فانه لابد أن يعطى الفاحص أو القائم بالبحث للمفحوصين اسبابا لأخذ الاختبار لها قيمة شخصية لهم ، وبدون هذا فان اداءهم او مشاركتهم فى الاختبار تكون عملا لا معنى له ، مما ينعكس على نوع الأداء وبالتالى على البيانات التى نقيم عليها البحث كله ··

(٢) الدوافع التى تخفض الدرجة : اما الدوافع التي تؤدى الى خفض الدرجة على الاختبار فنستطيع أن نحدد منها ما هو متعمد مقصود ، وما هو متصل بالرغبة الزائدة فى الأداء الممتاز والتفوق وما يصاحب هذا من قلق وتوتر ٠

وليس عجيبا أن يعمد مفحوص الى أن يكون اداؤه سيئا بحيث تكـون درجته على الاختبار أقل كثيرا مما يستطيع ، ولكنه يفعل هذا لأنه ، مثلا ، لا يريد أن يحرم من عضوية جماعة يحبها ويحرص على ان يبقى فيها وحصوله على درجة عالية فى الاختبار كفيل بأن يبعده عن جماعته الى جماعة اخرى أو لأنه يتوقع أن تضاف اليه أعباء وواجبات اضافية نتيجة حصوله على درجة عالية فى الاختبار وهكذا ٠

اما الدافع الشديد الى الامتياز والتفوق فى الأداء على اختبار ما فقد يؤدى الى التأثير السيىء على هذا الأداء ، كما أن التوتر الزائد يؤدى الى اخطاء فى الأداء كما نلاحظ فى الاختبارات اللغوية أو الاختبارات التى تتطلب دقة حركية وتآزرا حسيا حركيا بنوع خاص ٠

ومن المهم ان نشير الى أن قلق الاختبار ، كما اظهرت الأبحاث التى اجريت عنه ، متعلم من معاملة الآباء والمدرسين للتلاميذ ، كما انه متعلم من رفقاء التلاميذ واصحابهم ويبدو أن قلق الاختبار موجود بدرجات تكاد تتساوى بين القادرين من التلاميذ وبين غير القادرين منهم كما اثبتت أبحاث ساراسون Sarason وزملائه (٢٩) ، مستعملين فى قياس قلق الاختبار اختبارا وضعوه لهذا الغرض خاصة ٠

وقد وجدت بعض التجارب التى أجريت على قلق الاختبار علاقة سالبة بين القلق والأداء فى اختبارات للتحصيل واختبارات الذكاء ، مما يؤيد الفرض الذى يذهب الى الأطفال القلقين فى موقف الاختبار يهبط مستوى أدائهم ، ولكن ، كما تشير انستازى (١١ : ٣٦ ـ ٣٨) لا ينبغى ان ناخذ هذه العـلاقة السالبة بين القلق والأداء فى الاختبار على انها حتما اشارة الى علاقة سببية بينهما ، ذلك لأنه من المحتمل أن ينشأ عند المفحوص قلق لأن اداءه فى الاختبار

كان ضعيفا ، وليس العكس · ومما تشير اليه نتائج الدراسات فى هـذه الناحيـة ، ان درجـة بسـيطة من القـلق تسـاعد فى الأداء فى الاختبـارات وفى التعـلم ، بينمـا المسـتوى العـالى من القـلق يكون له اثر معطل · وتشير النتائج أيضا الى أن التعليمات الموجهة الى المفحوصين التى تتضمن توضيحا لأهمية الاختبار ونتائجه للمفحوص وتستدعى اندماجه فى الأداء فيه ، هذا النوع من التعليمات يحـدث اثرا ايجـابيا على اداء المفحوصين الذين يكون لديهم مستوى منخفض من القلق عندما يبداون الأداء فى الاختبار ، بينما تحدث أثرا سيئا عند اولئك المفحوصين الذين يبدأون بمستوى عال من القلق ·

على أن العلاقة بين القلق وبين الأداء فى الاختبارات ، وكذلك العلاقة بين القلق وبين التعلم ، تعتبر من المسائل التى تحتاج الى بحوث وتجـارب عديدة حتى تتحدد بصورة أدق وأوضح ، ويمكن الرجوع الى بعض نتـائج البحوث فى هذا الصدد الى بعض المراجع المتخصصة (٧) ·

ومن الطبيعى أن نتوقع أن كل موقف اختبارى يتضمن نوعا من التهديد، منه مثلا الخوف من الفشل ، أو الخوف من العقاب ، او الخوف من الاتهـام، أو من الاعتراف بقبول وجود سمة أو خاصية شـخصية قد يكشف عنهـا الاختبار ، مثل العصابية أو الميل الى التسلط · لهذا ، كان من المهم أن يؤكد الفاحص للمفحوص أن النتائج سوف تستخدم لصالحه ، مع تأكيد الجوانب الإيجابية بصفة خاصة من هذا الاستخدام ·

التهيئة لأخذ الاختبار :

(١) **ادراك المفحوص لمغزى الاختبار :** من المهم ان يرغب المفحـوص فى أن تكون درجته فى الاختبار صادقة صحيحة ، ذلك لأن هذه الرغبة ضمان لأن تكون الدرجة صادقة فعلا ، ويساعد التعاون بين الفاحص والمفحوص على ايجاد هذه الرغبة ، ولهذا كان امرا يستحق أن يحرص المختبر على توفيره، وهو ليس عسير التحقق على اى حال · على انه من المهم أن يكون واضحـا أن التعاون الذى نقصده بين المختبر والمفحوص ليس معناه ان تحال المسئولية كلها وتنقل الى المفحوص ، وانما نقصد أن يجعل الفاحص المفحوص يحس

باستمرار أنه مشارك له فى العمل مشاركة العضو العامل فى اى جماعة ، ولذا كان من الضرورى أن نتذكر أن الموقف الاختبارى المثالى هو الموقف الذى يكون المفحوص فيه شريكا فى اختبار نفسه ، كما سبقت الاشارة عند الكلام عن دوافع المفحوص ٠

وعلى الفاحص أن يعطى المفحوص كل ثقته فيما يتصل بهدف الاختبار، وأن يصور له الاختبار على أنه فرصة متاحة واداة علمية ميسرة يعرف عن طريقها شيئا عن نفسه ، تجعله أقدر على اتخاذ القرار المناسب له ، وعلى فهم ما يتخذ من قرارات متصلة به سواء فى العلاج أو الارشاد النفسى أو عند التقدم لكلية أو معهد أو وظيفة ما ٠ واذا كان الطبيب البارع يشرح لمريضه نوع الدواء الذى يعينه له والنتائج التى يتوقعها من تناوله الدواء ، فان المختبر الممتاز يجب أيضا أن يجعل الهدف من أخذ الاختبار والنتائج التى تترتب عليه واضحة بالنسبة للمفحوص ، ولا شك أن هذا الوضوح يجعل المفحوص اكثر حرصا على أن يجعل الاختبار يعطى صورة دقيقة بقدر الامكان عما يراد قياسـه ٠

(٢) **دور المعلومات التمهيدية** : يزداد الاعتراف فى الوقت الحاضر باهمية المعلومات السابقة التى يعطيها الفاحص للمفحـوصين فى تهيئتهم واعدادهم للاختبار ، ففى الارشاد النفسى مثلا (١٩) لم يعد اعطاء الاختبارات مسألة روتينية يقوم بها المفحوصون قبل أن تبدأ عملية الارشاد النفسى ، بل أصبح من الضرورى أن تسبق اعطاء الاختبارات مقابلات وغالبا ما يفضـل مقابلتان أو ثلاثة ، تتحدد فيها المشكلة ويناقش المختبر المفحوصين فى انسب الاختبارات للمشكلة والغرض الذى من أجله يسعى هؤلاء المفحوصون الى الارشاد النفسى ٠ هذا ، وان لمثل هذه المقابلات اثرا لا يمكن اغفاله وهو انها تساعد على الاقلال من قلق الاختبار الذى تحدثنا عنه ٠

(٣) **دور التدريب على الاختبار** : ظهـرت فى السنوات الأخيـرة برامج ومؤسسات لتدريب المفحوصين على أخذ الاختبارات النفسية ، كمـا ظهرت فى الأسواق سلاسل من « الاختبارات المصغرة » يتـدرب عليهـا المفحوصون قبل أن يتعرضوا لبرامج الانتقاء المهنى أو التعليمى ٠ وازدادت هذه الظاهرة شيوعا فى المجتمعات التى تلعب فيها الاختبارات دورا حاسما فى هذا المجال ، وخاصة الولايات المتحدة الأمريكية ٠

وقد شغلت مسألة التدريب على الاختبارات النفسية الباحثين فى ميدان القياس النفسى منذ نشأته المبكرة ، وأجريت تجارب هامة للتحقق من فعالية التدريب على الاختبارات ، ويمكن أن نلخص نتائج هذه البحوث فيما يلى (٢٠ : ٧٨ ـ ٨٠) ٠

(أ) تختلف برامج التدريب على الاختبارات فى طبيعتها ، وتعتمد فعاليتها على التهيؤ المبدئى لدى المفحوص وعلى دوافعه ٠ وقد لوحظ أن مجرد توافر الفرصة للمفحوص لأن يعطى الاختبار مرتين له بعض الآثر ، وخاصة اذا لم تكن لديه فى الأصل فكرة عن الاختبار ٠ كما لوحظ أيضا أن مجرد تخصيص عدد من الساعات لاعطاء اشارات وتلميحات عن طريقة اخذ الاختبار والتدريب عليه يمكن أن تزيد درجة المفحوص ما بين ٨ ، ١٠ نقاط ٠

(ب) تزداد فعالية التدريب كلما تحول من مجرد مراجعة سريعة الى مقرر منظم ٠ فعند تدريب بعض الطلاب على حل بعض المسائل الرياضية لمدة ٢١ ساعة من الدروس داخل الفصل و ٢١ ساعة أخرى من الواجبات المنزلية أدى ذلك الى زيادة الدرجة بمتوسط ١٦ نقطة فى أحد اختبارات الاستعداد الرياضى وحدث ما يشبه ذلك فى اختبارات الاستعداد اللغوى ، ويزداد متوسط عدد نقاط الدرجة بزيادة ساعات التدريب ٠ ومعنى ذلك أن التدريب على القدرات والمهارات اللازمة للاداء الاختبارى ، والتنبيه الى حيل الاجابة على الاختبارات النفسية له أثر مؤكد ٠

(حـ) يصدق قانون الغلة المتناقصة على نتائج التدريب على الاختبارات النفسية ٠ ومعنى ذلك أن زيادة نقاط درجات الاختبارات تزيد بزيادة ساعات التدريب حتى يصل عدد هذه الساعات الى نقطة معينة لا يترتب عليها أى زيادة مناسبة فى درجات الاختبارات ٠ فاذا كانت زيادة الدرجة بمقدار ١٠ نقاط فى الاختبار اللفظى تتطلب ١٢ ساعة من التدريب ؛ وفى الاختبار الرياضى تتطلب ٨ ساعات ، فان هذه الزيادة بمقدار ٤٠ نقطة تتطلب ١١٨٥ ساعة للاختبار اللفظى ، ١٠٧ ساعات للاختبار الرياضى ٠ ومعنى ذلك اننا نجد أن العائد ـ عند نقطة معينة ـ لا يتكافأ مع الجهد المبذول ٠

(د) يؤدى التدريب على الاختبارات النفسية الى ظهور أنواع جديدة

من الظلم الاجتماعى ، والى الخلل فى تكافؤ الفرص ، فلا شك ان القادرين وحدهم من حيث وفرة الوقت والمال هم الذين يتلقون هذه البرامج التدريبية مما يعينهم على اجتياز عوائق الانتقاء والالتحـــاق بالمؤسســات التعليمية (الكليات الجامعية مثلا) والمهنية التى يرغبون فيها ، بينما لا تتاح هذه الفرص لغير القادرين ٠

عملية الاختبار باعتبارها علاقة اجتماعية :

لا ينبغى أن يتصور الفاحص نفسه انسانا لا ينفعل ولا يتأثر ولا يحيـد عن التزامه بما تنص عليه واجباته وأعماله الرسمية فى الموقف الاختبـارى٠ وذلك لأن عملية الاختبار فى المجالات النفسية ليست عملية قياس « شىء » أو «موضوع» ليس بينه وبين الفاحص أية وشيجة، باستعمال أداة فنية دقيقة الى حد بعيد ، كما هو الحال فى قياس النواحى المادية التى يقوم بها العالم الطبيعى أو المهندس والتى يكون فيها الفاحص محايدا الى أقصى حد بالنسبة للموضوع المقاس ٠ أما والموضوع المقاس هو الانسان فان عملية الاختبـار لا يمكن أن تخلو من علاقات اجتماعية ونفسية على درجة كبيرة من التعقـد بين الفاحص وبين الشخص موضوع الاختبار ٠

(١) ادراك الذات وادراك الآخر عند الفاحص والمفحوص : المسألة فى العلاقة بين الفاحص والمفحوص أكبر وأعمق من مجرد أن تكون حرصا من جانب الفاحص على أن يشجع « المفحوص » وأن يكون « لطيفا » معه (١١ ، ٢١) كما هو شائع فى التوصيات التى توجه الى الفاحصين ٠ وانما الأمر يتضمن شخصين ، فى حالة الاختبار الفردى مثلا ، يمثل أحدهما ، وهـــو المفحوص ، شخصا عنده تصورات عن نفسه وعن الآخرين ، واقعية أو غيـر واقعية ، وتصورات عن الفاحص نفسه وعن الموقف الاختبارى ، وعنـــده كذلك مخاوف وشكوك وتوقعات وآمال ٠ وانه لحامل ذلك كله الى الموقف الاختبارى حيث يؤثر فيه من غير شك ٠ أما الشخص الآخر فهو الفاحص ، وهو انسان عنده أيضا توقعات وآمال ومخاوف وشكوك وان اختلفت نوعا ودرجة عما عند المفحوص ، بل وعنده تحيزات وتفضيلات لابد أن تنعكس وتؤثر فى كل مرحلة من مراحل الاختبار : اعطائه وتصحيحه وتفسيره ٠

ومن الطبيعى أنه عندما يأتى شخصان بهذه الحالة التى صـــورناها

موجزين الى موقف اختبار أو قياس أو تقويم فلا بد أن نتوقع نشأة علاقات اجتماعية ونفسية متشابكة تجعل مهمة المختبر اكثر دقة ورهافة ٠

وفيما يلى نذكر امثلة لهذا النوع من العلاقات تكفى لأن تلفت النظر الى امكان تنوع هذه العلاقات والى غموضها ٠ وتساعد الفاحص فى دور التدريب أو الممارسة الفعلية على أن ينمى عنده حساسية لهذه العلاقات فى المواقف الاختبارية ، بحيث يكتشفها بسرعة فيواجهها بصورة تجنبه خطأ أو تجاوزا يمكن تجنبه بالالتفات الى مثل ما سنشير اليه ، وتطبيقه فى مران طويل فى مواقف اختبارية تجريبية وحقيقية ٠

قد يحس شخص محول ليجرى عليه اختبار معين أو عدة اختبارات أن السلطة التى حولته تستبد به وتميل الى أن تسيطر عليه ، ويكون رد الفعل عنده هو التمرد على هذه السلطة ، وعلى الفاحص الذى يمثل « سلطة » أخرى بالنسبة اليه ٠ ويظهر هذا التمرد المنقول من السلطة الأصلية الى سلطة الفاحص فى مقاومات ومظاهر معينة أثناء الموقف الاختبارى ، كأن يعطى مثلا معلومات يعتقد أنها مضللة عن نفسه ، أو يعطى معلومات خاطئة عمدا ، أو يظهر اهمالا وعدم اكتراث فى الاجابة أو الأداء أيا كانت صورته ، أو يبدى اعتراضات كثيرة ، وما الى هذا من مظاهر المقاومة والفشل فى العلاقة بين المختبر والمفحوص التى تكلمنا عنها من قبل ٠ ولابد أن يدرك الفاحص هذا ، خاصة فى المواقف الفردية ، ويعمل على الاقلال من هذه المقاومة ومعالجتها بالاعتراف بوجودها وتقبلها وابرازها للمفحوص ، وتأكيد الضرر الذى سيلحق به من أثر هذه المقاومة ، ومظاهرها المعطلة على النتائج التى يعطيها الاختبار والتى هى فى النهاية لصالح المفحوص ، وبعد ذلك يمكن أن يعطى الاختبار ٠

بل ان الشخص الذى يأتى بنفسه طالبا اجراء اختبار أو عدة اختبارات عليه ، ربما لا يخلو هو أيضا من صعوبة على الرغم من توفر الدافع لأخذ الاختبار عنده ، ولو ظاهريا على الأقل ٠ والأغلب أن نتوقع تعاونا كبيرا بين مثل هذا الشخص وبين الفاحص ، ولكن قد تتعارض دوافع هذا المفحوص مع أهداف الاختبار ، فاذا كان الدافع عنده هو أن يتعرف على بعض جوانب من شخصيته ٠ فانه فى الوقت نفسه يخشى أن يكشف الاختبار ، أو الفاحص عن نواحى قصور عنده يحرص هو بكل وسيلة أن يبقيها خارج نطاق وعيه ٠

وأكثر من هذا ، فان مجرد سعى المرء الى اخذ اختبار ، او الى المساعدة النفسية ، فيه مساس باستقلاله الشخصى واحساسه بقيمته وقدرته بحيث يزيد هذا عنده القلق ازاء الاختبار الفاحص جميعا ٠

وقد يكون الدافع الظاهر ، أو المعلن ، للسعى لأخذ اختبار غير الدافع الحقيقى وراءه ، فالطالب الذى يريد أن يأخذ اختبارات لأغراض تعليمية أو للتوجيه المهنى قد يكتشف الفاحص المحنك أن وراء هذه الرغبة الظاهرة مشكلات نفسية واجتماعية هى التى حركت هذا الطالب الى طلب اخذ الاختبار وغالبا ما يكون هذا الاكتشاف بداية مرحلة من العلاج أو الارشاد النفسى قد يطلب بعده معلومات من اختبارات ، وانه لقمين ان يحسن استخدام هذه المعلومات بعد أن تحل المشكلات التى كانت محاولات ابعادها واخفائها تستنفد جهده وطاقته ، والتى كانت تعوق خطاه عن السير فى الاتجاه المناسب الذى يرتضيه ٠

ومن الممكن أن يغلب على المفحوص مظهر الخضوع الكامل للفاحص ، ويبدى من الرغبة فى التعاون أكثر كثيرا مما يقتضيه الموقف عادة ، ومثل هذا اذ يسلم نفسه للفاحص ، انما هو فى الواقع يعبر عن ميله الى التخلى عن مسئوليته عن نفسه بالقائها كاملة على الفاحص ، بل ان هذا الخضوع قد يخفى وراءه عدوانية أو مقاومة لابد أن يعالجها الفاحص ان كان ذلك فى مكنته او يحيلها الى من يستطيع ويوفق فيها ٠

اما عن الفاحص نفسه فانه ، كما قلنا ، يحمل بدوره الى الموقف الاختبارى تصورات عن نفسه وعن المفحوص ورغبات وشكوكا مما أشرنا اليه. وهذه قد تتوافق مع تلك التى عند المفحوص بحيث تدعم كل منها الأخرى ، وقد تتعارضان وتتضادان ٠ فاذا تصادف أن ذلك المفحوص ، الذى يلبس لباس الخضوع ويتخلى عن مسئوليته عن نفسه للفاحص ، وقع مع فاحص يميل الى السيطرة على الناس فان هذا الفاحص ، فى أغلب الظن ، لن يستطيع أن يكتشف أن هذه السلبية او الخضوع ليس الا حائلا يقيمه المفحوص ليحول بين المختبر وبين الوصول الى أعماق لا يريد له أن يصل اليها ٠ أما اذا تصادف وكان الفاحص من النوع الذى لا يرحب بأن تحال عليه تلك المسئولية، ربما بسبب [illegible] فى نفسه فانه لن يشعر بالا[illegible] فى التعامل مع المفحوص ،

وقد يلجأ الى الافراط فى استخدام الاختبارات وما تسفر عنه من نتائج لأن كثرة البيانات تعطيه احساسا أكبر من الأمن مع مثل ذلك المفحوص ·

وقد يكون عند الفاحص دوافع كامنة لا يتبينها اثناء قيامه بعملية الاختبار نفسها أو وراء اشتغاله بالاختبار بصفة عامة ، فقد يكون اختياره لهذا النوع من العمل اشباعا لحاجات معينة كأن يستشعر الكفاءة الاجتماعية من خلال التفاعل الاجتماعى مع الآخرين وراء أستار الاختبارات ورسمياتها لأنه يحس بعدم القدرة الاجتماعية فى مواقفها الطبيعية ، أو قد يكون كمن يقاوم شكوكا أو يجيب عن تساؤلات عن نفسه بمقارنة نتائج مفحوصيه بنفسه، او قد يكون فى تعامله مع المفحوصين ساعيا وراء السند العاطفى وذلك بأن يبدى رقة زائدة معهم ويكون عظيم التسامح والتساهل معهم ، مما قد يجعله يغفل جوانب ضعف فى نتائج الاختبار قد يكون لهـا مغزاها بالنسـبة له وللمفحوص على السواء (كالنقص اللغوى) · على أن هذا ليس معناه أن يتجنب الفاحص اى تعبير انفعالى مع المفحوص ، بل على العكس كما سبق ان اكدنا انه لا يصح أن يصطنع البرود أو الجمود أو الرسمية المفرطة ، وانما المهم هنا هو أن يكون الفاحص واعيا بدوافعه وحاجاته هو حتى يستطيع أن يبعد آثارها كلما استطاع عن التأثير فى الموقف الاختبارى الذى هو مسئول عنه ، على أن يحرص على اظهار التدعيم العاطفى المناسب الذى ييسر اقامة علاقة سليمة بينه وبين المفحوص ويقلل من مقاومته ويساعده لى استدعاء افضل وأصدق ما عنده فى أدائه فى الاختبار ·

(٢) دور اللغة فى تكوين العلاقة بين الفاحص والمفحوص : لعل القارىء يلاحظ أن معظم ما تضمنه موضوع ادراك كل من الفاحص والمفحوص لذاته وللآخر يدخل فى باب التواصل غير non-verbal communication بل ان بعضـه يدخـل فى مجـال « ما لا يمـكن التلفظ به » univerbalized ونشير هنا الى الدور المباشر للغة فى تكوين العلاقة بين الفاحص والمفحوص·

ولعل من الأمور الواضحة بذاتها انه اذا لم يفهـم كل من الفاحص والمفحوص لغة الآخر فان الموقف الاختبارى كله يدخل فى باب العبث أو منطق اللامعقول ، ولا نبالغ اذا قلنا انه يكون أدخل الى فن الملهاة ، وهذا كله بالطبع قد يكون مأساة بالنسبة للقرارات التى تعتمد على نتائج مثل هذه الاختبارات·

ويصدق هذا القول على الباحث الذى يحمل أدواته ويذهب بها الى احدى القبائل البدائية المنعزلة ثم يقدم لهم تعليمات وأسئلة بلغة لا يفهمها الا هو ، كما يصدق على الباحث الآخر الذى يختار عيناته من اعمار معينة (أطفال صغار مثلا) أو مستويات اقتصادية اجتماعية معينة ثم يطبق عليهم اختبارات لا يكاد يفهم تعليماتها او اسئلتها سواء ، بل قد لا يفهمها هو لأنها مترجمة ترجمة شبه حرفية عن لغة أخرى · نقول هذا القول لأنه يوجد فى المكتبة العربية فى الوقت الحاضر طوفان من الاختبارات لم يبذل فيها أصحابها جهدا الا محض الترجمة - التى قد تكون ركيكة - الى اللغة العربية، ونكاد نقول ان بعض هذه الاختبارات لا يكاد يفهمها لغويا اصحابها انفسهم·

بالطبع توجد لغة الايماءات والاشارات التى نجح علماء النفس فى استخدامها مع العاجزين لغويا ومع الذين لا يتحدثون اللغة القومية ومع الأطفال الذين لم يكتمل نموهم اللغوى بعد ، الا أن لهذه اللغة اصولها التى يجب اتقانها والتمكن منها ، والا أدخلتنا هى الأخرى فى فن العبث أو منطق اللامعقول ·

وقد اثبتت البحوث التى اجريت فى الولايات المتحدة (٢٠ : ٦٦ - ٦٧) - وهى مجتمع متعدد اللغات - أن الأطفال الذين يتقنون لغة معينة - كاللغة الأسبانية - يجب اختبارهم بهذه اللغة دون اللغة الانجليزية ، الا اذا كان الهدف من الاختبار قياس مدى اتقانهم اللغة الانجليزية · وبالطبع فاننا فى اختبار الأطفال يجب أن تستخدم لغة منتقاة من قاموسهم اللغوى، ومن الأفضل بالطبع - فى هذه الحالة وأمثالها كاختبار الأميين - استخدام اللغة العامية سواء فى اعطاء التعليمات أو فى صياغة أسئلة الاختبار · وفى هذا الصدد لابد للمتخصص فى مجال القياس النفسى أن يكون على صلة وثيقة بالباحثين فى مجال علم اللغة والمتخصصين فى اللهجات المحلية ، وأن يوفر لنفسه معاجم هذه اللهجات - ان وجدت · ومن المستحسن فى هذه الحالات أن يكون الفاحص الذى يطبق الاختبار على هذه العينات من المتحدثين الأصليين بهذه اللهجة · حتى لا ينشأ عن سوء الفهم أو سوء الاستخدام عائقا فى فهم المفحوصين للاختبار ·

الا ان اللغة لا تعمل فى جو اجتماعى فارغ أو أجوف ، فلابد فى جميع الأحوال من الدفء العاطفى وجو الصداقة الذى يوفره الفاحص للمفحوص

فى موقف الاختبار · وقد أثبتت البحوث التى اجريت على تحليل محتـــوى المقابلات الشخصية أنه حين تجرى مقابلات رسمية مع أطفـــــال من سن ٨ سنوات يعيشون فى حى الزنوج بمدينة نيويورك (حى هارلم) لم ينتج هؤلاء الأطفال ــ الا فيما ندر ــ استجابات تزيد عن كلمة واحدة ، سواء اجريت هذه المقابلات داخل المدرسة أو داخل المنزل · ولكن حين يستخدم الأسلوب غير الرسمى فى المقابلة ، كأن تتم فى حضور بعض أصدقاء الطفل والجلوس معهم حيث يجلسون ــ على الأرض مثلا ، ومشاركتهم فى شرب علب العصائر التى يشربونها ثم اجراء المقابلة فى صورة حديث عادى ، فان النتائج تثبت فى هذه الحالة أن الطفل الذى يحكم على لغته من خلال المواقف الرسمية للمقابلة بأنها غير ناضجة يندفع منطلقا فى حديث طلق صريح وغير متحفظ ·

ولعل القول الأخير يصدق أكثر ما يصدق على الاختبارات التى تتناول الجوانب الوجدانية والاجتماعية لشخصية المفحوص (الميول والاتجـــاهات مثلا) · ولعل أشهر الأمثلة على ذلك تطبيق مقاييس الاتجاهات التى تتناول المسائل الاجتماعية الخلافية الحساسة ــ كالاتجاه نحو تحديد النسل فى مصر ــ ان اجابات المفحوصين على هذه المقاييس تأتى مباشرة وقصيرة وتتفق مع قواعد « المرغوبية الاجتماعية » اذا لجأ الفاحص الى الأسلوب الرسمى ، ولكن اذا تبسط مع المفحوصين ، كأن يجلس مع الفلاحين على « المصطبة » ، ويشرب معهم قهوتهم أو شايهم ، فان هذا اللون من « الأحاديث العادية » يطلق الألسنة من عقالها ، ويجعل الموقف الاختبارى أكثر طبيعية ، وحينئذ تؤدى اللغــــة دورها الطبيعى فى تكوين العلاقة الطيبة بين الفاحص والمفحوص ·

كيف يشوه الفاحصون نتائج الاختبارات ؟

يسجل تاريخ العلم امثلة على خداع الذات الذى قد يقع فيه باحثون وملاحظون مدربون · وكثير من هذه الأخطاء يحقق لهؤلاء الفاحصين نوعا من الخدمة الذاتية ، قد لا يكون من بينها المكافآت المادية المباشرة لهم ، وانما قد يكون الأهم عن ذلك اشباع حاجات الفاحص الى تدعيم آرائه ومعتقداته ·

وبالطبع فان تشويه نتائج الاختبارات أثناء عملية التقويم النفسى خطأ بالغ ، ان لم يكن خطرا عظيما · لنفرض أن أحد الباحثين التربويين صمم

طريقة جديدة فى تدريس الفهم اللغوى ويرغب ـ ولو لاشعوريا ـ فى ان يثبت كفاءتها وفعاليتها ٠ انه قد يطبق الصورة (ا) من اختبار الفهم اللغوى قبليا ، والصورة (ب) بعديا ، وبالطبع فانه فى هذه الحالة قد يضمن الحصول على نتائج ايجابية لصـــالح طريقة التدريس المقتـــرحة اذا ذكر للمفحوصين اثناء القياس القبلى أن درجات الاختبار ليست لها أهمية ثم قرأ عليهم تعليمات الاختبار بسرعة ، وربما يسىء قراءة زمن ساعته فيختصر زمن العمل فى الاختبار ٠ أما اذا استخدم الفاحص عكس ذلك فانه قـد يرفع من درجات مفحوصى المجموعة الضابطة اثناء القياس القبلى ، أو درجات مفحوصى المجموعة التجريبية أثناء القياس البعدى ٠ وفى الحالتين يحصل على نتائج مختلفة قد يساء فهمها وتأويلها ٠

وبعض أخطاء الفاحصين قد تكون مقصودة ـ وهذه يمكن ان تكون من قبيل الجرائم العلمية التى يذكر تاريخ العلم قليلا منها وصل بعضها بالباحثين الى حد تزييف الحقائق وفبركة النتائج (٨) ٠ اما « الأخطاء التى تحركها الدوافع » motivated error فقد تكون غير مقصودة ٠ فاى اجـــراء مقنن لتطبيق الاختبار ، مهما بلغت دقته لا يمنع حـــدوث بعض الآثار غير المقصودة فى درجات الاختبار حين يتوقع المعلم مثلا من تلميذه « النجيب » أن يؤدى أداء جيداً فيه، وحين يضغط تحيز الفاحص عليه ضغطا لا شعوريا فيظهر الرغبة فى أن يحرز المفحوص من جنسه أو دينه أو بيئته أو طبقته الاجتماعية تفوقا فى الاختبار ٠ ويعود هذا مرة أخرى الى حاجة المفحوص الى معـرفة « ذاته » وفهم « شخصيته » من حيث ميوله واتجاهاته وقيمه وتفضيلاته وتحيزاته وارائه وأفكاره ٠ وكلما ازداد استبصار الفاحص بذاته كان اكثر « موضوعية » فى مواجهة مواقف الاختبار ٠

الا أن هذا الاجراء يحتاج الى جهد منظم وتدريب شاق ، ولعلنا نعيد تنظيم طرق تدريس مقررات الاختبارات النفسية والقياس النفسى والتقـــويم التربوى والنفسى فى الجامعات بحيث تجعل مسالة تدريب الفاحصين بجوانبها المختلفة التى تناولناها فى هذا الفصل محور برامج هذه المقررات ٠ ويمكننا ان نستفيد أيضاً من بعض ما يستخدمه الباحثون فى مجالات اخرى ـ كالطب ـ حيث يستخدم أســلوب « الفحص الأعمى المـــزدوج » double-blind لتجنب الوقوع فى خداع الذات ٠ ومن ذلك أن يعطى لنصف المفحوصين عقار

محايد لا يختلف على أى نحو ظاهرى عن الدواء الفعلى بحيث لا يستطيع الطبيب أو المريض أو مساعد العمل التمييز بين المجموعتين ، ولا يتم هـذا التمييز الا أثناء تحليل النتائج تبعا لنظام ترميزهما • وبالطبع فان هذا الاجراء لا يزال صعب التنفيذ فى مجال البحوث النفسية والتربوية • والبديل هو ما يقترحه بعض الثقاة فى الميدان مثل كرونباك (٢١ : ٦٨) ـ من ضرورة أن تطبق الاختبارات فى حالة استخدام المجموعات التجريبية والضابطة بواسـطة فاحصين من الخارج لا يعرفون مقدما انواع المجموعات • ثم قبل تصحيح الاختبارات وتفسير نتائجها أن تخلط جميع الأوراق ـ الاختبارات القبلية والبعدية للمجموعات التجريبية والضابطة ـ خلطا عشوائيا ، مع الاحتفـاظ بمفتاح للتعرف على هذه الأوراق فيما بعد ، لا يكون متاحا للملاحظين أو المصححين •

وتوجد عوامل مباشرة تكمن وراء تشويه نتائج الاختبارات ، سواء كان مقصودا أو غير مقصود • ومن ذلك ما يترتب على درجات الاختبارات من نفع أو خسارة يتعرض لهما الفاحص ، ومن ذلك المعلم الذى تعتمد ترقيته أو مكافاته على نتائج فصله فقد يؤثر فى أداء تلاميذه فى الاختبارات بالغش المباشر أو التـلويح به ، ويذكـر هـورن Horne وجـارتى Garty فى بحثهما عام ١٩٨١ أمثلة كثيرة لما يحدث فى هذا الصدد فى المـدارس الأمريكية (٢١ : ٦٨) •

الفصـل الرابع

ثبات الاختبـارات

معنى ثبات الاختبار :

يقصد بمصطلح الثبات reliability في علم القياس النفسي دقة الاختبار في القياس أو الملاحظة وعدم تناقصه مع نفسه ، واتساقه واطراده فيما يزودنا به من معلومات عن سلوك المفحوص • ولابد للباحث أن يحـدد درجة الموثوقية في بيانات الاختبار كما تتمثل في اتســاق القيـاسات أو الملاحظات التي يحصل عليها • ويمكن التحقق من هذه الموثوقية بصفة عامة باحدى وسيلتين : اما بتكرار تطبيق الاختبار نفسه على نفس المفحوصين ، أو باستخدام اختبارات مماثلة •

وبالطبع فان من حقائق القياس النفسي أن درجات الاختبارات تختلف من موقف لآخر ، وقد يرجع ذلك الى اختلاف درجة الانتباه عند المفحوص ومقدار جهده المبذول من وقت لآخر ، وذلك بالنسبة للفترات الزمنية القصيرة، أما اذا طالت الفترة فان التغيرات في درجات الاختبارات قد ترجع الى أو التعلم أو التغيرات في الجوانب الصحية أو في شخصية المفحوص • فاذا استخدمت اختبارات مختلفة ـ ولكن متماثلة ـ فان الاختلاف في الدرجـات قد يرجع الى مدى ألفة المفحوص بالأسئلة في كل من الحالتين •

الا أن أخطر العوامل التي قد تؤدي الى اختلاف الدرجات في المواقف الاختبارية المختلفة بالنسبة للاختبار الواحد ولنفس المفحوصين ، هو عامل الصـــدفة chance ، وهو عامل يرجع الى الظروف العشوائية التي لا يتحكم فيها الفاحص مثل تخمين المفحوص للاجابة أو التذبذبات والتقلبات الوقتية في انتباهه ، ويطلق على هذا العامل خطأ القياس error of measurement •

وفي عملية التقويم النفسي لا يمكن أن يعتمد قرار الفاحص على أخطاء المصادفة هذه ، وخاصة اذا كان القرار مصيريا مثل انتقاء الشخص لمهنة

أو دراسة معينة ، أو اعطائه شهادة أو اجازة بالكفاءة أو الخبرة ، أو انهاء العلاج بالنسبة لمريض ، أو الحاق طفل متخلف عقليا باحدى المؤسسات ٠

ويستخدم مفهوم الثبات بمعناه الواسع ليدل على مدى اعتماد الفروق الفردية فى درجات الاختبار على أخطاء الصدفة المتضمنة فى القياس (والتى تدل على عدم الثبات) أو على الاختلافات الحقيقية الأصيلة فى الصفة أو الخاصية السيكولوجية موضع الاهتمام (والتى تدل على الصدق كما سيتبين فى الفصل التالى) ٠ ويمكن صياغة المشكلة صياغة فنية على النحو الآتى :

اذا افترضنا ان الدرجات التى يحصل عليها المفحوصين فى اختبار معين تمثل التباين الكلى total variance يصبح السؤال ماهى النسبة فى هذا التباين الكلى التى يمكن اعتبارها من نوع تباين الخطأ error variance وتلك التى ترجع الى التباين الحقيقى ؟ true variance ان الاجابة على هذا السؤال ليست سهلة كما يبدو لأن ما نعتبره أحيانا من نوع تباين الخطأ فى موقف معين يمكن اعتباره تباينا حقيقيا فى موقف آخر ٠ فمثلا اذا كان اهتمامنا متجها نحو دراسة التقلب الوجـــدانى فان التغيرات اليومية فى درجات اختبار يقيس « الانشراح والانقباض » تصبح مرتبطة بغرض الاختبار نفسه وتكون بالتالى جزءا من التباين الحقيقى فى الدرجات ٠ أما اذا كان الاختبار يهدف الى قياس الخصائص الأكثر استقرارا فى الشخصية كالذكاء أو الميل العلمى تصبح هذه التغيرات اليومية ذاتها من نوع تباين الخطأ ٠

ويمكن التعبير عن انواع التباين المشار اليها بالمعادلة الآتية :

التباين الكلى = التباين الحقيقى + تباين الخطأ ٠

ومن هذه المعادلة نستنتج أنه اذا زاد تباين الخطأ يقل التباين الحقيقى ويؤدى ذلك الى تذبذب الاختبار فيما يقيس وعدم اتساقه أو بعبارة أخرى عدم ثباته ، والعكس صحيح أى اذا زاد التباين الحقيقى يقل تباين الخطأ ويؤدى ذلك الى مزيد من استقرار الاختبار وثباته ٠

وعلى ذلك يمكننا ان نعرف تباين الخطأ بأنه يتعلق بالشروط والظروف التى لا تتصل من قريب أو بعيد بغرض الاختبار ٠ ولذلك فحينما يحاول المدرس

أو المرشد النفسى أن يقنن الاختبار بأن يحافظ على توحيد الظروف الاختبارية وذلك بضبط مكان الاختبار وتعليماته وزمنه وعلاقة بالمختبرين وغيرها من العوامل التى تحدثنا عنها فى الفصل السابق فانه بذلك يقلل من فرص حدوث تباين الخطأ ويجعل درجة الاختبار أكثر ثباتا واستقرارا . ومع ذلك فلايمكن اعتبار الاختبار اداة ثابتة ثباتا كاملا يصل الى درجة ثبات مقاييس الظواهر الطبيعية (كالمتر للاطوال والكيلو جرام للاوزان) . ولذلك لابد ان يحدد صانعو الاختبارات مقدار ثباتهافى كراسة التعليمات بحيث يكون دليلا للاخصائيين النفسيين حينما تطبق هذه الاختبارات فى ظروف معيارية مقننة .

والمشكلة الجوهرية فى تحديد ثبات الاجراءات الاختبارية بالنسبة لمجموعة معينة من الأفراد هى فى التمييز بين التباين الحقيقى والتباين الذى يرجع الى الخطأ ، وعلاقة ذلك بنوع الاستنتاج الذى يرغب الباحث النفسى فى التوصل اليه من درجات الاختبار . وبعدما يتم هذا التمييز بين نوعى التباين يبدأ الباحث فى تصميم سلسلة من العمليات الامبريقية (التجريبية) والاحصائية تفيد فى الوصول الى أفضل تقدير لنسبة كل من نوعى التباين .

مصادر التباين :

١ - **الخصائص المستمرة والعامة فى الفرد :** فالتباين فى درجات ثورنديك أول من أعد تصنيفا شاملا لهذه المصادر عام ١٩٥١ فى فصل عظيم كتبه عن « الثبات » فى كتاب لنديكويست الشهير (٢٦) ، وقد عدل ستانلى J.C. Stanley هذا التصنيف فى فصل جديد بنفس العنوان فى الطبعة الثانية من هذا الكتاب والتى حررها روبرت ثورنديك (٣١) ، وتشمل هذه القائمة المعدلة ما يلى (٣١ : ٣٦٤) :

١ - **الخصائص المستمرة والعامة فى الفرد :** فالتباين فى درجات الأفراد ينشأ فى العادة نتيجة لما يظهره الأفراد المختلفون من مقادير مختلفة فى سمات أو خصائص معينة ذات طبيعة مستمرة وعامة ، ومن ذلك الاختلاف فى مستوى قدرة معينة لها دورها فى الأداء فى عدد كبير من الاختبارات ويؤدى هذا الاختلاف الى فروق فردية ظاهرة ومن أمثلة ذلك القدرة على الاستدلال أو السهولة العددية أو الفهم اللغوى . وعلى هذا توجد فى الأفراد

الاختلافات بقدر ما يوجد من سمات أو قدرات ، وهى كثيرة العـــدد بالغة التنوع لا يتسع المقام لحصرها ،ويمكن الرجوع اليها فى المراجع المتخصصة(٩)٠ ويمكن القول أن أى أداء اختبارى يعتمد جزئيا على القدرات والسمات التى تتداخل بالفعل فى عدد من الاختبارات النفسية ٠

ولعل القارىء يلاحظ ان هذا المصدر من مصادر التباين يمثل خصائص مستمرة ومستقرة لدى الفرد ، ويؤدى بالضرورة الى فروق فردية مستقرة وحقيقية فى الأداء الاختبارى ، وهو نوع من التباين المنتظم على نحو يعين الباحث على تناوله فى أى اجراء يتخذه عند تقدير الثبات ٠

ويضيف ستانلى الى فئة الخصائص المستمرة والعامة مصدرين اخرين للتباين هما : القدرة العامة على فهم التعليمات والخبرة بأخذ الاختبارات test-wisenes وهما عاملان يكادان يؤثران فى كل درجة اختبارية سواء اراد الفاحص أو لم يرد ٠ ويلعب عامل فهم التعليمات دورا هاما على وجه الخصوص حين يكون الاختبار جديدا ، وكانت تعليماته معقدة ٠ اما العامل الثانى فيتمثل فيما يظهره المفحوص « الخبير بالاختبار » من خصائص تختلف عن المفحوص الأقل خبرة ، فالأول يكون عادة اقل انفعالية ، وأمهر فى التخمين ، واكثر وعيا بالمنبهات الثانوية والخارجية التى تفيد فى الاجابة على أسئلة الاختبارات ٠ ولعلك تلاحظ ان وجـــود تباين ناتج عن هذين المصدرين يعد من الأمور غير المرغوب فيها فى ضوء الغرض من الاختبـــار الا اذا ارتبطا بقدرة عقلية أساسية ٠ ويمكن التقليل من اثرهما بصفة عامة عن طريق تحسين تعليمات الاختبار من ناحية وتقديم قدر من التدريب على أسئلة الاختبار من ناحية أخرى ٠

وتمثل هذه الفئة من مصادر التباين تحديا حقيقيا لمؤلفى الاختبارات ، لأنها تتطلب التحكم ، الا اذا كان وجودها مطلوبا ومرغوبا فيه وفى هذه الحالةتصبح منوع التباين الحقيقى٠ ويتطلب هذا من الباحث النفسى ان يقدم لمفحوصيه تعليمات على اعلى درجة من الوضـــوح وباقل قدر من المنبهات الثانوية ٠

٢ ـ **الخصائص المستمرة والخاصة فى الفرد** : قد ينشأ التباين فى الدرجات الاختبارية نتيجة لخصائص معينة فى الافراد تتسم بالاستمرار من

ناحية الا انها تكون على درجة من الخصوصية للمجال النوعى الذى نختبره، ومن أمثلة ذلك ما نجده فى اختبارات التهجى او القسمة المطولة دون سواها من الاختبارات • فبالاضافة الى التباين الذى يميز الأداء العام فى التهجى أو الحساب يوجد نوع اخر من التباين يرتبط بصورة معينة من صور الاختبار أو طريقة معينة فى هذا الاختبار كما هو الحال مثلا فى سلوك التهجى حين يكون على المفحوص أن يتهجى كلمات معينة شفويا او يكتبها صحيحة فى الاملاء أو يتعرف على الأخطاء فى بعض الكلمات المطبوعة • واختبار القسمة المطولة الذىيتطلب انتاج الاستجابة أو اختيار الاجابة من عدة بدائل (الاختيار من متعدد) • وقد ينشأ عن ذلك نوع من التباين بالطبع • ويوجد مصدر آخر للتباين ينتمى الى هذه الفئة هو طبيعة عينة الأسئلة أو المفردات التى يتألف منها الاختبار وينشأ عن الفروق الفردية فىتفاصيل المعارف والمهارات التى لدى الأفراد المختلفين • وعلى هذا لو أعد الباحث اختبـــارين من عينتين من الأسئلة اختيرنا مستقلتين من نفس الأصل الاحصائى للاسئلة او المفردات فقد لا يحصل الفرد الواحد على نفس الدرجة فيهما بسبب التنوع بين الأفراد فى المهارات والمعلومات اللازمة للاجابة على اسئلة كل من الاختبارين • وفى هذه الحالة تنشأ الحاجة الى المقارنة بين الأفراد فى اختبارات متماثلة فيما تقيس وتختلف فى انواع المفردات أو طريقة العرض للتحكم فى هذا المصدر باعتباره ينتمى الى تباين الخطأ •

٣ - الخصائص العامة المؤقتة فى الفرد او الموقف الاختبارى : ومن أمثلة هذه الخصائص الصحة الجسمية ومقدار النوم الذى حصل عليه الفرد فى الليلة السابقة على الاختبار ، والتعب ، وحالة الدافعية ، والتوتر الانفعالى وغيرها من العوامل الداخلية فى الفرد والتى تؤثر على ادائه فى الاختبار ، بالاضافة الى العوامل الخارجية فى الموقف الاختبارى كدرجة حرارة الجو ، والاضاءة والتهوية وغيرها • ومن المعلوم ان أداء الفـــرد فى الاختبارات المختلفة يتأثر بهذه العوامل بدرجات مختلفة ، الا ان جميع الاختبارات تتأثر بها على نحو او اخر • كما تختلف هذه العوامل فى درجـــة استمرارها وعموميتها ، فبعضها قد يتغير من يوم لآخر ، وبعضها الآخـــر من ساعة لأخرى ، وبعضها الثالث من لحظة لأخرى • ويمكن لنا أن نتصور هـــذه العوامل على انها تميز الفرد فى جلسة اختبار معينة دون سواها • ويحتاج الباحث النفسى فى مثل الأحوال الى التحقق من اتساق الأداء فى الاختبار

سواء فى لحظة معينة هى لحظة تطبيق الاختبـار ، أو من فترة زمنية الى أخرى (بعد عدة أيام أو عدة أسابيع) كما أن على الباحث أن يحدد ما اذا كانت هذه الخصائص المؤقتة من المصادر المنتظمة الحقيقية للتباين أم هى من مصادر تباين الخطأ ·

٤ ـ **الخصائص الخاصة المؤقتة فى الفرد** : ويشمل هذه الفئة العوامل ذات النطاق الضيق سواء من حيث الزمن أو المدى اذا قورنت بعوامل الفئـة السابقة ، وبعضها يرتبط بالاختبار ككل ، وبعضها يرتبط بأسئلته ، ومن أمثلة النوع الأول ما يتصل بصعوبة التعليمات فى حالة الاختبارات الجديدة التى قد تنشأ عنها فروق فردية فى فهم طبيعة المهام التى يتألف منها الاختبار، وقد يؤدى ذلك الى تباين عشوائى ينتمى الى تباين الخطأ · وكذلك قد تتطلب الاجابة على الاختبار استراتيجيات مختلفة يستخدمها الأفراد المختلفون ، بعضها قد يكون أكفأ من البعض الآخر ، وحينئذ ينشأ مصدر جديد لتباين الخطأ · وبالاضافة الى ذلك فان الأداء فى بعض الاختبارات ـ وخاصـة اختبارات المهارات الحركية ـ قد يكون عرضة للتحسن نتيجة للممارسـة ، ويؤثر فى ذلك مستوى الممارسة لدى الفرد عند بدء الاختبار · كما قد توجد مجموعة من العوامل المتنوعة التى يشملها مصطلح « التأهب لأداء الاختبار» مثل التركيز على السرعة أو الدقة فى اختبارات القدرات الموقوتة ، أو الوعى ببعض المنبهات فى حالة الاختبارات الادراكية ، أو الحالة المزاجيـة لدى الفرد فى حالة مقاييس الاتجاهات أو قوائم الميول · وهذه العوامل ـ على الرغم من أهميتها ـ يمكن تجاهلها فى الاختبارات البسيطة والمقننة ، الا أنها تلعب دورا أكبر فى الاختبارات الموقوتة ، وفى الاختبارات التى تتكون من أنواع جديدة من الأسئلة أو المفردات ، وفى المقاييس التى تتطلب مقدارا كبيرا من التفسير الاستبطانى ، وفى الأداء الاختبارى الذى لم يتجـاوز المرحلة التكوينية (أى لم يصل فيه المفحوص الى اجابة نهـائية ، وانما تكـون اجابته على هيئة مسودة مبدئية) ·

وتوجد مجموعة أخرى من العوامل تنتمى الى هـذا الفئة وتخص المفردات أو الأسئلة ذاتها التى يتكون منها الاختبار ، أو بضـع لحظـات من الزمن الكلى للاختبار وتتمثل فى التذبذبات الوقتية فى الذاكرة أو الانتباه،

او التعويق المؤقت للافراد ، او التنوع الدورى فى مقدار الجهد المبذول ، وهى جميعا تنتمى الى تباين الخطأ •

٥ - العوامل المنتظمة او العشوائية التى تؤثر فى تطبيق الاختبار او تصحيحه : قد ينشأ التباين من مصادر اخرى غير المفحوص ذاته وتنشأ عن ظروف اعطاء الاختبار وتصحيحه ، ومن ذلك زمن الاختبار وتعليماته ومقدار الضوضاء وتشتت الانتباه وفشل الاضاءة ونقص مواد الاختبار (كالأقلام) وغيرها ، وهى جميعا تختلف من شخص لآخر أو من مجموعة لأخرى وهذا النوع من الاختلاف ينشأ على وجه الخصوص حين يطبق الاختبار الواحد فاحصون مختلفون ، أو فى الاختبارات التى تحتاج الى توقيت دقيق او تلك التى تتطلب تعليمات معقدة للغاية • كما قد ينشأ هذا المصدر للتباين عند تصحيح أو تقدير الدرجة فى الاختبار وخاصة حين يعتمد ذلك على حكم الفاحص او الملاحظ أو المصحح ، ومن ذلك اعطاء الدرجة او الحكم فى اختبارات المقال ، ومقاييس التقدير او الأساليب الاسقاطية او فى اى اختبار يتطلب قدرا من التفسير أو الحكم من جانب الفاحص - اننا فى هذه الحالة قد نجدنا ازاء نوع من التباين يرجع فى أصله الى اختلاف الفاحصين أو الملاحظين أو المصححين ، أو الى التغيرات التى تطرأ على الفاحص الواحد من وقت لآخر ، أو الى تناقض لا يخضع للتحكم أو التنبؤ داخل الفاحص أو المصحح نفسه • وهذه العوامل جميعا تنتمى الى تباين الخطأ •

٦ - العوامل التى ترجع الى الصدفة البحتة : يوجد نوع آخر من التباين لا تستطيع أن تعود به الى عامل بذاته (أو عوامل بذاتها) ولذلك يمكن أن نعتبر عوامله من النوع العشوائى أو الذى يرجع الى الصدفة ، ومن ذلك تخمين المفحوص للاجابة فى الاختبار حيث يلعب « الحظ » دورا واضحا فى انتقاء الاجابة (والتى قد تكون صحيحة) ، كما قد يرجع الى عوامل أخرى كثيرة غامضة تلعب دورها فى تشتيت انتباه المفحوص فى لحظة معينة وهذه العوامل تنتمى الى تباين الخطأ بمعناه المباشر •

طرق حساب الثبات :

توجد بالطبع انواع من ثبات الاختبار بعدد مصادر التباين التى اشرنا

اليها ، اى بعدد الظروف التى يمكن أن تؤثر فى درجات هـــذه الاختبارات ولا تكون كلها او بعضها متعلقة بغرض الاختبار ، وتوجد طرق تستخدم فى حساب الثبات يشير كل منها الى المصادر التى يمكن أن يرجع اليها تباين الخطأ • وهذه الطرق ليست بدائل يحل بعضها محل بعض ، وانما لكل منها استخدامه الخاص حسب مصدر تباين الخطأ الذى نهتم به •

١ ـ اعادة الاختبار :

اشرنا الى أن من المصادر الأساسية لتباين الخطأ ما يتمثل فى التغيرات العشوائية التى تنتاب الأفراد من جلسة اختبار الى اخرى • وقد تنتج هذه التقلبات عن ظروف تطرأ على الموقف الاختبارى ولا يمكن التحكم فيها مثل التغيرات الفجائية فى الجو أو حدوث ضوضاء مفاجئة (وغيرها من مشتتات الانتباه) • وقد تنتج عن تغيرات فى ظروف المفحوص نفسه كما تتمثل فى حالات المرض والتعب والاضطراب الانفعالى والقلق وغيرها • وهذه العوامل تدل على المدى الذى تتأثر به درجات الاختبار بالتقلبات اليومية العشوائية فى المفحوص أو فى ظروف الموقف الاختبارى •

ومن الواضح أن درجة الاستقرار الوقتى Temporal Stability تعتمد الى حد ما على الفترة الزمنية التى تنقضى على تطبيق الاختبار والتى نقدر بعدها ثبات الاختبار • وتوجد أمثلة عديدة لاختبارات لها ثبات مرتفع اذا قدر هذا الثبات بعد فترات زمنية قصيرة (أيام قلائل او اسابيع قلائل) الا ان ثباتها ينخفض انخفاضا شديدا عندما تمتد هـــذه الفترات الى عدة سنوات • وكثير من اختبارات ذكاء اطفال ما قبل المدرسة لها ثبات مرتفع خلال فترة ما قبل المدرسة ولكنها تفشل فى التنبؤ بذكاء الأطفال فى مرحلة الطفولة المتأخرة أو المراهقة أو النضج بسبب ما يطرأ على الطفل من تغيرات ارتقائية فى النواحى العقلية •

ولا شك فى أن التذبذبات ذات المدى الزمنى القصير يمكن اعتبارها من نوع تباين الخطأ فى درجة الاختبار • ولذلك ففى الممارسة العملية عندما تقوم بتقدير مثل هذا النوع من الثبات باعادة الاختبار لابد من ان تكون الفترة الزمنية بين مرتى التطبيق قصيرة ، وتكون أقصر عند اختبارنا للاطفـــال

الصغار منها عند اختبارنا للكبار بسبب سرعة معدل النمو فى الطفولة عنه فى مراحل النمو الأخرى · وبالنسبة لأى نوع من المفحوصين فان قواعد الخبرة تؤكد انه يجب الا تزيد فترة ما بين الاختبارين عن ستة اشهر ·

هذا التحفظ مهم للغاية ، لأننا نعلم جميعا أن السلوك الانسانى فى حالة تغير مستمر ، فاذا كانت هذه التغيرات تحدث فى الأداء الاختبارى للمفحوصين بعد انقضاء فترات زمنية طويلة فانها يمكن أن تعد ضربا من النمو أو التدهور فى السلوك لا مجرد تغيرات عشوائية ، بل قد تنتمى هذه التغيرات الى ميدان من السلوك أوسع نطاقا مما يتناوله الأداء الاختبارى ذاته ، فالمفحوص قد يتغير أداؤه فى اختبارات الاستعداد المدرسى او الفهم الميكانيكى أو الحكم الفنى بعد عشر سنوات بسبب تعرضه لخبرات غير عادية خلال هذه الفترة · بعبارة أخرى يمكن للفـــرد أن يتغير مركزه النسبى فى الاختبار بعد هذه الفترة صعودا أو هبوطا بالنسبة لأقرانه من نفس العمر بسبب ظروف خاصة تعود الى المنزل أو المدرسة او البيئـــة الاجتماعية أو بسبب المرض المزمن أو الاضطراب الانفعالى المستمر ·

ويجب أن نشير الى أن العمليات النفسية ذاتها قد تختلف فيما بينها فى مدى ما تظهره من تذبذبات وتقلبات وقتية · فمثلا قد يتأثر اتزان الحركات الدقيقة للاصابع بأى تغير ولو كان طفيفا فى ظروف المفحوص او فى ظروف الموقف الاختبارى بينما نجد الاستدلال لا يسهل التأثير فيـــه بالمؤثرات العارضة · ولذلك اذا أردنا أن نحصل على تقدير دقيق لثبات اختبار اتزان الأصابع يجب اعادة تطبيقه مرات عدة فى قيام متوالية ، بينما قد تكفى جلسة اختبار واحدة لقياس الاستدلال دون حاجة الى اعادة التطبيق · ومعنى ذلك أنه من الضرورى فهم طبيعة السلوك الذى يقيسه الاختبار قبل اختيار انسب الطرق لحساب ثباته ·

ويقدر معامل الثبات الذى يتأثر بتباين الخطـــأ الناتج عن التقلبات الوقتية بطريقة اعادة الاختبار test-retest . ويطبق الاختبار فى المـــرة الثانية ، بصورته ذاتها على نفس الأفراد ، وقد يتطلب ذلك اعادة التطبيق عدة مرات فى مناسبات مختلفة ، ويحسب معــــامل الارتباط بين درجات المفحوصين فى مرتى التطبيق أو متوسط معاملات الارتباط اذا أعيد تطبيق

الاختبار عدة مرات ، ولعلك تلاحظ أن درجات الأفراد قد تختلف فى المرة (أو المرات) المختلفة ، وهذا لا يهم كثيرا انما المهم ان اوضاعهم النسبية لا تتغير كثيرا ، ومعامل الارتباط فى جوهره هو طريقة تحديد الوضع النسبى ويسمى المعامل الناتج معامل الاستقرار coefficient of stability .

ورغم بساطة هذه الطريقة فانها تتضمن بعض المشكلات · منها مثلا أن الممارسة قد تنتج قدرا من التحسن فى درجات اعادة الاختبار · ومنها أيضا أنه لو كانت الفترة بين موقف الاختبار واعادته قصيرة فان المفحوص قد يستدعى بعض استجاباته السابقة ، أو بعبارة أخرى قد نحصل على نفس النمط من الاجابة من حيث الوضع النسبى للصواب والخطأ ولكن ليس بسبب العملية النفسية موضع القياس ، وبالتالى لا تكون الدرجات التى نحصل عليها من موقفى الاختبار مستقلة بعضها عن بعض ، كما ان معامل الارتباط بين مجموعتى النتائج قد يكون مرتفعا ارتفاعا زائفا · وكذلك من مشكلات طريقة إعادة الاختبار أن طبيعة الاختبار ذاتها تتغير بسبب الاعادة وخاصة اذا كانت أسئلة الاختبار تتطلب استخدام بعض العمليات العليا كالاستدلال أو الابتكار · ففى مثل هذه الأسئلة نجد أن المفحوص اذا استطاع اكتشاف العلاقة أو أعطى إستجابة جديدة أو توصل الى حل المشكلة ، فانه يمكن أن يستعيد اجابته ذاتها فى المستقبل اعتمادا على الذاكرة دون التفكير ، ومعنى ذلك أن طبيعة الأداء فى الموقفين تكون مختلفة · ومع ذلك فان معامل الارتباط بين مرتى التطبيق قد يكون مرتفعا ارتفاعا زائفا ، ومع ذلك لا يدل على ثبات حقيقى · وعلى ذلك فان هذه الطريقة لا تصلح الا مع الاختبارات التى لا تتأثر تأثيرا كبيرا بالاعادة أو التكرار مثل اختبارات التمييز الحسى السمعى أو البصرى · اما بالنسبة لمعظم الاختبارات النفسية فان طريقة اعادة الاختبار بالرغم من شيوعها فى الاستخدام قد لا تكون ملائمة لها ·

٢ ـ الصور المتكافئة او البديلة :

من المألوف لدى كثير من الطلاب أن يدخلوا الامتحان « فيسعدهم الحظ» بعدد كبير من الأسئلة يتناول الموضوعات التى اهتموا بدراستها ومذاكرتها اكثر من غيرها ، وقد يحدث العكس فى كثير من الأحيان هذا الموقف يدل على مصدر اخر من مصادر تباين الخطأ فى درجات الاختبار وهو مدى اعتماد

درجة المفحوص فى الاختبار على عوامل تتعلق بطريقة اختيار الأسئلة . فاذا افتراضنا أن قام باحث آخر باعداد اختبار تتوافر فيه نفس مواصفات الاختبار الأول مع اختلاف فى المحتوى فالى أى حـــد تختلف درجة المفحـــوص فى الاختبارين ؟ الواقع أن الفرق فى الدرجة التى يحصل عليها نفس الأفراد فى هذين الاختبارين هو نوع الثبات التى يتأثر بعينة الأسئلة التى يتكون منها كل من بينهما . فمثلا اذا افترضنا أن باحثا قام باعداد اختبـــار لقياس المعلومات الميكانيكية يتكون من ٥٠ سؤالا ، ثم قام باحث آخر باعداد اختبار آخر مكافىء له (أى يتكون من ٥٠ سؤالا مختلفا تقيس نفس العملية النفسية)، وانفرض أن أسئلة كل من الاختبارين قد صيغت بنفس الدرجـــة من المهارة بحيث يتناول نفس المستوى من الصعوبة ، فانه تبعا لعوامل الصدفة البحتة الناجمة عن الفروق الفردية فى الخبرات السابقة للمفحوصين قـــد تختلف الصعوبة النسبية لمجموعتى الأسئلة من شخص لآخر ، وعلى ذلك فان الاختبار الأول قد يحتوى على عدد كبير من الأسئلة لم يألفه المفحوص (أ) بينما تكون الفته بأسئلة الاختبار الثانى أكبر . وقـــد يحدث العـــكس بالنســـبة للمفحوص (ب) .

فاذا افترضنا أن هذين المفحوصين متســـاويان تقريبا فى معلوماتهم الميكانيكية (أى فى درجاتهم الحقيقية أو التباين الحقيقى) فانَ المفحوص (ب) سوف يتفوق على المفحوص (أ) فى الاختبار الأول ، بينما سوف يتفوق (أ) على (ب) فى الاختبار الثـــانى . ومعنى ذلك أن الموقف النســـبى للمفحوصين قد يتغير بل قد ينعكس ويمثل هذا نوعا من تباين الخطأ يرجع الى الصدفة الناتجة من طبيعة محتوى كل من صورتى الاختبار .

ويقدر معامل الثبات فى هذه الحالة بتطبيق صورتى الاختبار فى جلسة واحدة أو فى جلستين مختلفتين على نفس الأفراد ، ثم يحسب معامل الارتباط من الدرجات فى الاختبارين ويسمى معامل الثبات فى هذه الحالة بمعامل التكافؤ Coefficient of equivalence . والمهم فى جميع الحالات ان تكون صور الاختبار متكافئة فعلا ، أى تقيسان نفس العمليات النفسية بنفس المستوى من الصعوبة أو الألفة ، وأن تؤثر هذه العمليات النفســـية فى الأداء فى الاختبار بنفس الطريقة . ولكن ، كما سبق ان اوضحنا ، قد تؤثر الصفات الخاصة تأثيرا مختلفا فى الاختبارين ومن ذلك المعلومات الخاصة التى قد

تساعد المفحوص فى الاجابة على سؤال معين من اسئلة احـد الاختبارين ولكنها لا تساعده فى الاختبار الآخر ٠

وتبـــاين الخطأ المتضمن هنـا يــدل على التغيـــرات فى الأداء من مجموعة معينة من الأسئلة الى مجموعة أخرى ، اى الخطأ الذى يرجع الى عينة الأسئلة ولكنه لا يتضمن تباين الخطأ الذى يعود الى العوامل الوقتية ، والذى سبق أن تحدثنا عنه ، والذى يتمثل فى التغيرات فى الأداء من موقف اختبارى الى آخر ٠ ويمكن الجمع بين هذين النوعين من تباين الخطأ اذا قمنا بتطبيق الصورتين المتكافئتين من الاختبار فى مناســبتين مختلفتين (اى لا نطبقهما فى نفس الوقت) ٠ ويمكن بهذه الطريقة تفادى بعض مشـكلات طريقة اعاد الاختبار ذاتها ٠ وفى هذه الحالة يدل معامل الثبات على كل من الاستقرار والتكافؤ معا ، ويسمى بهذا الاسم ٠

وعند اعداد الصور المكافئة equivalent او البـــديلة alternate من الاختبارات يجب أن نتأكد كما قلنا من أن هذه الصور فعـــلا متكافئة ٠ والصور المكافئة هى فى جوهرها اختبارات تم اعداد كل منها على حـــدة وبطريقة مستقلة بحيث تتوافر فيها نفس المواصفات ، اى أن تحتوى على نفس العدد من الأسئلة ٠ وأن تكون صياغة هذه الأسئلة متماثلة ، وأن تتضمن الأسئلة محتوى متشابها ، وأن تتعادل الأسئلة فى مستويات الصعوبة ، وأن تتفق الاختبارات فى جميع المظاهر الأخرى مثل التعليمات والزمن والأمثلة التوضيحية والشكل العام ٠ ولا يمكن أن تعزى الفـــروق فى الدرجات من اختبار الى آخر الى تباين الخطأ الذى يرجع الى عينة محتوى الاختبار الا اذا كانت الصورتان متكافئتين أو متعادلتين حقا ٠

ومن المؤكد أن حاجتنا الى الصور المتكافئة من الاختبارات تتعدى مجرد تحديد ثبات الاختبار ٠ فالصور البديلة تفيد كثيرا فى الدراسات التى نجريها حول آثار عامل تجريبى معين فى الأداء الاختبارى وخاصة اذا تطلب الأمر قياسا قبليا وبعديا ، كما يفيد استخدام الصور البديلة فى اعطاء وسـيلة لخفض احتمال التدريب أو الغش ٠ الا أن أهم صعوبات هذه الطريقة تتمثل فى عدم امكان اعداد صور متكافئة لكثير من الاختبارات لأن ذلك يضاعف الجهد والوقت والتكلفة المبذولة فى بناء الاختبار ٠ ومن المعروف أن اعداد

اختبار نفسى واحد يتطلب جهدا عظيما ، فما بالك باعداد اكثر من اختبار لقياس نفس السمة ٠

٣ ـ التجزئة النصفية :

يدل تجانس أسئلة الاختبار على مدى اتساق واطراد أداء المفحوص على جميع الأسئلة التى يتكون منها الاختبار ٠ فاذا كان لدينا اختباران تتكون أسئلة الأول منهما من عمليات ضرب فقط بينما تتكون اسئلة الاختبار الثانى من العمليات الحسابية الأساسية الأربعة فان الأداء فى الاختبار الأول سوف يكون أكثر اتساقا واطرادا بين أسئلته من الاختبار الثانى ، لأنه فى الاختبار الثانى (وهو اختبار غير متجانس الأسئلة) قد يحسن بعض المفحوصين الأداء فى عمليات الطرح مثلا بينما يسيئون أداء باقى العمليات ٠ فاذا زاد عدم تجانس الأسئلة ، كأن يتكون الاختبار مثلا من ٤٠ سؤالا يقيس كل عشرة منها عمليات نفسية مختلفة كالعلاقات المكانية والاستدلال الحسابى والمفردات والسرعة الادراكية فانه قد تكون العلاقة ضـــئيلة للغاية (ان كانت هناك علاقة) بين أداء المفحوص على هذه الأنماط الأربعة المختلفة من الأسئلة ٠

من ذلك نتبين ان درجة الشخص فى اختبار ما تكون أكثر وضـــوحا عندما نحصل عليها من اختبارات متجانسة البنية نسبيا ٠ وتبرز مشـــكلة وضوح معنى الدرجة الاختبارية اذا عدنا مرة أخرى الى المثال الأخير ، فاذا حصل المفحوصان (أ) و (ب) على درجة مقدارها ٢٠ فى هذا الاختبـــار فهل يمكننا ان نستنتج من ذلك أن أداءهما متساو تماما فى هذا الاختبار ؟ الاجابة بالنفى دون شك ٠ فقد ينجح المفحوص (أ) مثلا فى الاجابة على اسئلة المفردات والسرعة الادراكية ويفشل تماما فى الاجـــابة على اسئلة الاستدلال الحسابى والعلاقات المكانية ٠ أما المفحوص (ب) فقد يحصل على درجته من الاجابة على ٥ اسئلة للسرعة الادراكية و ٥ اسئلة اخرى للعلاقات المكانية و ١٠ اسئلة للاستدلال الحسابى ويفشل تماما فى الاجابة على اسئلة المفردات ٠ ويوجد عدد كبير من الاحتمالات الأخرى للدرجات التى يحصل عليها المفحوصون من الاختبارات غير المتجانسة فى البنية ٠ ومعنى ذلك أن درجة المفحوص فى مثل هذا الاختبار لها معان مختلفة ، بينما يكون معناها أكثر وضوحا فى الاختبارات المتجانسة ٠

والدرجة فى الاختبار المتجانس الأسئلة (معانى المفردات مثلا) قد لا يكون لها ايضا معنى واحد تماما فى جميع الحالات ، وانما قد تعنى مثلا أن المفحوص أجاب اجابة صحيحة على المفردات العشرين الأولى تقريبا اذا كانت هذه المفردات مرتبة حسب درجة الصعوبة ، وربما قد يترك كلمتين أو ثلاث من الكلمات السهلة أو قد يفشل فى الاجابة عليها ، وقد ينجح فى الاجابة على كلمة أو كلمتين أصعب وأبعد فى الترتيب ومع ذلك فان هذه الاختلافات الفردية قليلة وهينة اذا قورنت بتلك التى نلاحظها فى الاختبارات الأقل تجانسا ٠

وهكذا يعد عدم تجانس أسئلة الاختبار مصدرا لتباين الخطأ ، وبالتالى يؤثر فى ثبات الاختبار ٠ ولكن هل يختلف تجانس الأسئلة كمصدر لتباين الخطأ عن المصدر الذى سبق الحديث عنه وهو ملاءمة عينة المحتوى عموما ؟ الواقع أن الفرق بين هذين المصدرين يمكن أن يتضح اذا اعطينا مثالا متطرفا، لنفترض أن هناك اختبارا يقيس كل سؤال فيه عملية نفسية مختلفة ولا يرتبط ادنى ارتباط بالأسئلة الأخرى التى يقيس كل منها عمليات اخرى ٠ اننا فى مثل هذه الحالة يمكن أن نعد صورة متكافئة من هذا الاختبار تحتوى على نفس انواع الأسئلة وتوزيعها على نفس النحو الذى هى عليه فى الصورة الأولى ٠ وعندما نعطى الصورتين المتكافئتين قد تتفق درجاتهما اتفاقا كاملا مما يدل من الوجهة النظرية على معامل ثبات مرتفع لا يتأثر تأثرا كبيرا بعينة الأسئلة كمصدر لتباين الخطأ ٠ ولكن عندما نهتم بالثبات فى ضوء تجانس الأسئلة نجد ان هذا المعامل يصل الى الصفر أو يقترب منه ٠ حيث أن اطراد الأداء من سؤال الى آخر فى أى من صورتى الاختبار قد لا يتجاوز حدود الصدفة ٠

وللتحقق من اثر عدم تجانس الأسئلة كمصدر لتباين الخطأ فى الاختبار يلجأ الباحث الى تطبيق صورة واحدة من الاختبار مرة واحدة ، ثم يقوم عند تصحيح الاختبار بتجزئته الى نصفين وعندئذ يحصل على درجتين لكل مفحوص كل منهما فى كل نصف من نصفى الاختبار يمكنه المقارنة بينهما ويسمى هذه الطريقة التجزئة النصفية split-half ٠ ويحسب الباحث معامل الارتباط بين درجات الاختبار التى حصل عليها المفحوصون فى نصفى الاختبار، ويسمى معامل الثبات الناتج معامل الاتساق الداخلى consistency

internal ، حيث أن الأمر لا يحتاج اكثر من تطبيق صورة واحدة للاختبار مرة واحدة فقط ، وعملية التجزئة تتم أثناء تصحيح الاختبار وليس أثناء تطبيقه ·

ولعل أخطر المشكلات التى تواجهها طريقة التجزئة النصفية هى كيف نجزىء الاختبار للحصول على أفضل قسمين للمقارنة ، اى قسمين متكافئين، وخاصة اذا علمنا أن معظم الاختبارات لايمكن تجزئتها الى نصفين صالحين للمقارنة بينهما بمجرد تتبع التسلسل الرقمى للاسئلة ، وذلك بسبب الفروق فى طبيعة الأسئلة ومستوى صعوبتها بالاضافة الى الآثار التراكمية للحماس فى الأداء أو العمل والممارسة والتعب والملل وغيرها من العوامل التى تختلف فى بداية الاختبـــار عنها فى نهايته · ولذلك فان احـــدى الطرق الدقيقة والموضوعية للحصول على قسمين متناظرين ومتكافئين فى للاختبار أن يكون الاختبار فى اعداده المبدئى قد تحددت فيه مستويات الصعوبة لكل سؤال وذلك بايجاد النسب المئوية للافراد الذين أجابوا عليه اجابة صحيحة ، وعندئذ توزع الأسئلة على قسمى الاختبار تبعا لتكافؤ مستوى الصعوبة وتشابه المحتوى · فاذا كان ترتيب الأسئلة فى الاختبار متسلسلا تبعا لمستويات السهولة او الصعوبة فان الباحث قد يستخدم طريقة صالحة لمعظم الأغـــراض وفى نفس الوقت لا تستغرق وقتا طويلا ولا تتطلب جهدا زائدا وهى تجزئة الاختبار الى نصفين احدهما يضم الأسئلة ذات الأرقام الزوجية والآخر يضم الأسئلة ذات الأرقام الفردية ، الا اننا يجب أن نحذر أنه فى حالات مجموعات الأسئلة التى تتعلق بمشكلة واحدة (مثل الأسئلة التى تتلو فقرة معينة فى اختبار لفهم قراءة) فان هذه الفقرة باسئلتها جميعا يجب أن توضع فى احد قسمى الاختبار من الاختبار ، لأننا لو وزعنا مجموعة الأسئلة التى من هذا النوع على نصفى الاختبار فان التشابه بينهما سوف يزداد زيادة مصطنعة او زائفة ، وأى خطأ أو صواب فى فهم الفقرة قد يؤثر فى الاجابة على نصفى الاختبار ، ويؤدى ذلك الى ارتفاع معامل الارتباط بين النصفين دون أن يدل فى واقع الأمر على ثبات مرتفع ·

لقد بينا اننا نحصل على معامل الاتساق الداخلى بحســـاب معامل الارتباط بين درجات نصفى الاختبار الا أننا يجب أن نشير الى ان مثل هذا

المعامل يعطينا ثباتا لنصف الاختبار فقط وليس للاختبار كله • وهذه المشكلة لا توجد بالطبع فى الطريقتين السابقتين (اعادة الاختبار والصور المتكافئة) •

ونحن نعلم أنه اذا تساوت الظروف الأخرى فان الاختبار الأطول فى عدد أسئلته يكون أعلى ثباتا ، لأننا لو حصلنا على عينة اكبر من السلوك نصل الى مقاييس اكثر دقة وأكثر استقرارا • ويمكن تقدير الأثر الذى تحدثه اطالة الاختبار أو اختصاره على معامل الثبات باستخدام المعادلة التنبؤية الشهيرة التى اقترحها سبيرمان وبراون Spearman-Brown وهى :

$$ر = \frac{ن ر_١}{١ + (ن - ١) ر_١}$$

حيث أن ر = معامل الثبات الذى نريد الحصول عليه

$ر_١$ = معامل الثبات الذى حصلنا عليه

ن = عدد مرات اطالة الاختبار أو اختصاره اى انه اذا زاد عدد الأسئلة من ٢٥ الى ١٠٠ سؤال فان ن تكون ٤ أما اذا اختصر من ٦٠ الى ٣٠ سؤال فانها تصبح $\frac{١}{٢}$ •

ويمكن اختصار المعادلة حينما تطبق على معامل الثبات المقدر بطريقة التجزئة النصفية لأنها تتضمن اطالة الاختبار مرتين ، فتصبح على النحو الآتى !

$$ر = \frac{٢ ر_١}{١ + ر_١}$$

الا ان معادلة سبيرمان - براون تفترض أن التغيرات فى درجات نصفى الاختبار متساوية تماما أى تفترض تكافؤ ثباتهما ، وهو افتراض يصعب فى كثير من الأحوال توافره حتى ولو بدا نصفا الاختبار على درجة كبيرة من

التكافؤ بالفعل · وقد اقترح رولون Rulon تبسيط معادلة سبيرمان ـ براون مع وضع تباين نصفى الاختبار فى الاعتبار ، وصاغ لذلك المعادلة الآتية:

$$ر = ١ - \left[\frac{ع^٢ف}{ع^٢ك}\right]$$

حيث ع^٢ف = تباين فروق درجات نصفى الاختبار ·

ع^٢ك = تباين الاختبار ككل ·

وقد يستعاض عن ذلك باستخدام معادلة جتمان Guttman التى تستبعد هذا الافتراض وتصلح لحساب معامل الثبات مباشرة دون حاجة الى حساب معامل الارتباط · ويمكن صياغة هذه المعادلة على النحو الآتى :

$$ر = ٢ \left[١ - \left(\frac{ع^٢_١ + ع^٢_٢}{ع^٢ك} \right) \right]$$

وفى هذه المعادلة الرمز $ع^٢_١$ ، $ع^٢_٢$ = تباين النصف (١) والنصف (٢) فى الاختبار

ع^٢ك = تباين الاختبار ككل ·

٤ ـ طريقة كيودور ـ ريتشارد سون :

تصلح طريقة التجزئةالنصفية لحساب معمل الاتساق الداخلى (الاتساق بين نصفى الاختبار) ، أما بالنسبة لموضوع تجانس الأسئلة ذاتها والتى تؤلف فى مجموعها بنية الاختبار فان بعض العلماء يرون أنه لا يدخل فى نطاق مناقشة موضوع ثبات الاختبارات وانما هو من خصائص الاختبارات الجيدة بصفة عامة الا أن هذا الرأى لا يتفق مع التصور العام لمفهوم تباين الخطأ الذى عرضناه فما سبق ، فنحن ، دون شك ، فى حاجة الى معرفة مدى استقرار استجابات المفحوص على مفردات الاختبار سؤالا بعد آخر وهذا

ما يسمى « الاتساق بين أسئلة الاختبار » interitem consistency ، وهو مصطلح آخر قد لا يعنى نفس ما يعنيه مصطلح معامل الاتساق الداخلى فى طريقة التجزئة النصفية ، وحتى لا يحدث الخلط بين الطريقتين نسمى مصدر التباين هنا بتجانس الاختبار homogeneity ، وهو مصدر هام لأن عدم تجانس أسئلة الاختبار وتنوعها الشديد قد يؤدى الى اختلاف أداء المفحوصين وتغيره فى المواقف المختلفة ، فقد يوجد لدينا اختباران لهما معاملات ثبات مرتفعةحسبت اما بالصور المتكافئة أو بالتجزئة النصفيةولكنهما قد يختلفان فى المعاملات الخاصة بدرجة تجانس الأسئلة فيهما ، وقد تظهر هذه المشكلة فى صورتى الاختبار أو نصفيه ٠ ٠ والواقع أن الفرق بين معامل الاتساق الداخلى (الذى يحسب بالتجزئة النصفية كما سبق أن أوضحنا) وبين معامل التجانس الذى نشير اليه هنا يمكن اعتباره دليلا على مدى التجانس أو التباين فى محتوى الاختبار ٠

واشهر الطرق لحساب معامل التجانس طريقة كيودر - ريتشارد سون وهذه الطريقة تتشابه مع طريقة التجزئة النصفية فى انها تستخدم اجراء اختباريا واحدا لصورة واحدة من الاختبار ، ولكنها بدلا من أن تتطلب تجزئة الاختبار الى نصفين متكافئين فانها تعتمد على فحص أداء المفحوص على كل سؤال على حدة ، واشهر المعادلات الاحصائية التى تستخدم فى هذا الصدد هى :

$$ر = \left(\frac{ن}{ن - ١} \right) \frac{ع^{٢} - مج\,ص\,خ}{ع^{٢}}$$

حيث أن ر = معامل ثبات الاختبار

ن = عدد الأسئلة التى يتكون منها الاختبار ٠

$ع^{٢}$ = تباين الاختبار ككل ٠

مج ص خ = يحسب باعداد نسبة الأفراد الذين أصابوا فى اجابتهم على كل سؤال ونسبة الذين أخطأوا فيه ثم نضرب

النسبتين فى بعضهما ونجمع حواصل الضرب بالنسبة لجميع الأسئلة ٠

ويؤكد كرونباك (٢٠) أن معامل التجانس كما يحسب بطريقة كيودر ـ ريتشاردسون هو فى الواقع متوسط معاملات الاتساق الداخلى الناتجة عن تجزئات متعددة للاختبار ٠ اما معامل الاتساق الداخلى فيعتمد على تجزئة تهدف الى الحصول فقط على مجموعتين متساويتين او متكافئتين من الأسئلة، والتى تعطى معاملا للثبات بالرغم من عدم تجانس الاختبار ، ولعل من الأمثلة المتطرقة على ذلك أن يكون كل سؤال زوجى وفردى فى الاختبار يقيس عملية نفسية مختلفة عن أى سؤال فردى وزوجى آخر فى الاختبار ٠

ومن الواضح ان معادلة كيودر ـ ريتشاردسون تصلح للاختبارات التى تتم الاجابة على اسئلتها بالثنائية ، من نوع الكل أو لا شىء مثل صواب ـ خطأ ، نعم ـ لا ، موافق ـ معارض وغيرها ، ولهذا يقترح احد مؤلفى هـذا الكتاب (٩) أن يحل الرمزان (ا) ، (ب) محل (ص) ، (خ) فى المعادلة الأصلية لتصبح أوسع فى الاستخدام وتمتد الى نطاق اختبارات الاداء المميز٠

٥ ـ معامل الفا لكرونباك :

كثير من الاختبارات النفسية لا يقوم على نظام الكل أو لا شىء المشار اليه عند الحديث عن طريقة كيودر ـ ريتشاردسون ٠ ففى التقرير الذاتى قد يحصل المفحوص على درجة مختلفة تبعا لدرجة اجابته على سؤال يطلب تقدير مدى حدوث سلوك معين بانه يحدث « عادة » أو « احيانا » او « قليلا » أو « نادرا » ، أو عند الاجابة على سؤال فى مقياس للاتجاهات بقوله ازاء عبارة معينة بانه « موافق جدا » أو « موافق » أو « لا رأى له » أو « معارض » أو « معارض جدا » ٠ وبالطبع فان هذا النوع من الاختبارات قد يكون اكثر حاجة الى تحديد معامل التجانس فيه أكثر من اختبارات الأداء الأقصى التى هى فى العادة من نوع الكل أو لا شىء ٠ ولهذا اقترح كرونباك منذ عـام ١٩٥١ معادلة تطورت بعد ذلك على يد كايزر وميشيل عام ١٩٧٥ اطلق عليها معامل الفا Coefficient Alpha ، وفيها يحل مجموع تباينات درجات الأسئلة محل مجـ ص خ (أو مجـ ا ب اذا شئت) ٠ وصيغت المعادلة كما يلى :

$$\text{ر} = \left(\frac{\text{ن}}{\text{ن} - ١}\right)\left(\frac{\text{ع}^{٢} - \text{مج}\,(\text{ع}^{٢}_{١،\,٢،\,\text{ن}})}{\text{ع}^{٢}}\right)$$

أ

حيث الرمز مج ع^{٢} $_{١، ٢، ٣، ..، ن}$ = مجموع تباينات الأسئلة من السؤال الأول فى الاختبار الى السؤال الأخير ٠

٦ ـ ثبات الفاحصين والمصححين :

بعد أن نستبعد الأنواع الثلاثة السابقة من تباين الخطأ من التباين الكلى فان التباين المتبقى قد يرجع الى الفروق الفردية الحقيقية موضوع الاهتمام ، وبعضه الآخر الى عوامل أخرى غير مرتبطة بالاختبار ولكن يمكن ضبطها تجريبيا ٠ فمثلا لا يسجل الباحثون أخطاء القياس الناتجة عن تطبيق الاختبار فى ظروف مشتتة للانتباه ٠ أو فى زمن أطول أو أقصر مما يحدده كتيب الاختبار ، أو اعطاء تعليمات مختلفة عن التعليمات المحددة ، لأن هذه الأخطاء جميعا يمكن استبعادها والتغلب عليها فى الموقف الاختبارى نفسه اذا أتقن الفاحص فن اعطاء الاختبارات الذى عرضناه فى الفصل الثالث٠ ولذلك فنحن لا نجد فى العادة معاملات ثبات خاصة تشـــير الى « التباين الخاص بتشتت الانتباه » أو « تباين اختلاف الزمن » أو « تباين اختلاف التعليمات » الخ ٠ فمعظم الاختبارات تحدد اجراءات مقننة للتطبيق (أشرنا اليها فى الفصل الثالث) والتصحيح (سنشير اليها فى الفصل السادس) بحيث نفترض مع وجود هذه الاجراءات والالتزام بها أن « ثبات الفاحص » و « ثبات المصحح » يصل الى درجة كبيرة من الكفاءة تجعل الاختبار مفيدا فى الأغراض التطبيقية والعملية وبالتالى لا نكون فى حاجة الى تقدير مثل هذا النوع من الثبات الذى يرتبط فى جوهره بموضـــوعية الملاحظـــة أو التصحيح ٠ ويصدق هذا القول على وجه الخصوص بالنســـبة للاختبارات الموضوعية الجماعية التى نطبقها على جمـــاعات كبيرة والتى تصححها الحاسبات الالكترونية ٠ ففى هذه الحالات جميعا لا يتطلب الأمر الا التأكد من أن الفاحص يتتبع اجراءات التطبيق والتصحيح تتبعا دقيقا وتكون المشكلة فى جوهرها الضبط المحكم لظروف اجراء الاختبار ٠

الا أننا نلاحظ أنه فى الاختبارات الفردية مثل ستانفرد ـ بينيه يلعب المنحن دورا أكثر تعقيدا · لأن اجراءات الاختبارات فى هذه الأحوال ليست مقننة تقنينا وأمدا بل يعتمد معظمها على كفاءة تدريب الفاحص ، وعلى نجاحه فى تكوين علاقة بينه وبين المفحوص ، وعلى دوافع المفحوص · وقد يتطلب الأمر أن يقوم الفاحص بتقويم أداء المفحوص خلال عملية تطبيق الاختبار لأن مثل هذا الأداء يحدد كيف يستمر الفاحص فى موقف الاختبار · ولذلك فقد يختلف الفاحصون فى النتائج التى يحصلون عليها من نفس المفحوصين حتى ولو توافرت فى هؤلاء الفاحصين أفضل الشروط · وهذا الاختلاف قد يكون من نوع تباين الخطأ الذى يمكن أن ترجعه الى الفروق الفردية بين الفاحصين ويؤثر بالطبع فى التباين الكلى لدرجات المفحوصين · ولذلك فاذا كانت الاختبارات التى نستخدمها من النوع الذى تلعب فيه فردية الفاحص دورا هاما ، أو اذا كانت من النوع الذى يتأثر بالعوامل الذاتية وطرق الملاحظة المختلفة ، نكون فى حاجة ماسة الى تقدير « ثبات الفاحص » وخاصة اذا كان من الضرورى معرفة العلاقة بين تقديرات أكثر من فاحص لنفس المفحوص أو المفحوصين ·

وبالمثل نجد أن بعض الاختبارات تثير مشكلة « ثبات المصحح » وخاصة اختبارات المقال والاختبارات الاسقاطية بوجه عام ، وهى الأدوات التى تسمح بمقدار كبير من التفسير الذاتى من جانب المصحح وهو الفاحص أيضا فى كثير من الأحوال ·

ومن الواضح أن معامل ثبات المصحح يمكن الحصول عليه بحساب معامل الارتباط بين الدرجات التى يعطيها مصححان أو أكثر لنفس الأفراد فى نفس الاختبارات · ومعنى ذلك أن كل مفحوص يحصل على درجتين أو أكثر من تصحيح اختبار واحد · وهذه الطريقة لا تتضمن بالطبع نواحى القصور الموجودة فى طريقة اعادة الاختبار · أما معامل ثبات الفاحص فهو أكثر صعوبة لأنه بطبيعته يتضمن اعادة اختبار من جانب أكثر من فاحص ، وبالتالى فان قيام أكثر من فاحص باعادة الاختبار يعنى الوقوع فى جوانب القصور التى سبق الحديث عنها فى طريقة اعادة الاختبار التى يقوم بها نفس الفاحصين · بالاضافة الى أن معامل الارتباط الناتج يعكس عوامل الاستقرار الوقتى وثبات المصحح معا · ويمكن اللجوء الى نمط مختلف من التخطيط

التجريبى يعتمد على تحليل التباين لعزل تباين الخطأ الذى يمكن ارجاعه الى الفاحصين • وعلى كل فيمكن ان نعتبر كلا من معامل ثبات المصحح ومعامل ثبات الفاحص من نوع واحـــــد ونطلق عليه اسم معامل الموضــــوعية Coefficient of objectivity .

ويمكن ان نلخص أنواع ماممعلات الثبات وطرق تقديرها فى الجدول رقم (١) •

جدول (١) أنواع تباين الخطأ وعـــلاقته بطرق حساب الثبات (عن ١١ : ١١٨) •

المعامل	طريقة تقديره	تباين الخطأ
معامل الاستقرار	اعادة الاختبار بنفس صــورته على نفس المفحوصـــين فى مناسبة مختلفة	عينة الوقت
معامل التكافؤ	تطبيق الاختبار مع صـــورة مكافئة له فى نفس الوقت	عينة المحتوى
معامل الاستقرار والتكافؤ معا	تطبيق الاختبار ثم تطبيق صورة مكافئة له فى مناسبة مختلفة	عينــة الوقت ، عينة المحتوى
معامل الاتساق الداخلى	التجزئة النصفية	عينة المحتوى ، تجانس المحتوى
معامل التجانس	تحليل المفردات بطريقة كيودر - ريتشاردسون أو بطريقة الفا	عينة المحتوى تجانس المحتوى
معامل الموضوعية	تكرار المصحح أو الفاحص ذاتية المصـحح والفاحص عينة الوقت •	ذاتية الفاحص عينة الوقت

العوامل المؤثرة فى ثبات الاختبار :

١ - درجة اعتماد الاختبار على الزمن والسرعة :

تمثل الاختبارات ذات الزمن المحدد مشكلة من نوع خاص فى موضوع

الثبات ، وخاصة اذا كانت الفروق الفردية فى درجات هذه الاختبارات تعتمد على سرعة الأداء • ولذلك لا يمكن استخدام طريقتى التجزئة النصـــفية أو كيودر - ريتشارد سون لحساب ثبات هذه الاختبارات لأن هاتين الطريقتين تعتمدان على الاطراد فى ارتكاب الأخطاء • فمثلا اذا كان الاختبار يتكون من ٥٠ سؤالا من النوع الذى يعتمد على السرعة اعتمادا كليا (أى على عدد الأسئلة التى يحاولها المفحوص فى جلســة الاختبــار) كاختبــار النقر أو اختبار العمليات العددية لطلاب الثانوية العامة مثلا فاننا نجد ان المفحوص الذى يحصل على الدرجة ٣٦ مثلا تكون درجته فى كل من النصف الفردى والنصف الزوجى من الاختبار ١٨ ، بعبارة أخرى تتساوى تقريبا درجته فى نصفى الاختبار الا اذا تأثرت هذه الدرجة بالأخطاء العارضة التى ترجع الى اهماله لعدد قليل من هذه الأسئلة • وفيما عدا ذلك فان الارتباط بين الدرجات الفردية والدرجات الزوجية يكون أقرب الى الارتباط التام ، وهو معامل زائف لا يدل بحال من الأحوال على ثبات مرتفع الاختبار •

ولذلك فأفضل طرق حساب الثبات لهذه الاختبارات هى طريقتا اعادة الاختبار واستخدام الصور المتكافئة • ويمكن اللجوء الى طريقة التجزئة النصفية بشرط أن تتم التجزئة على أساس السرعة فى الأداء لا على أساس عدد الأخطاء • بعبارة أخرى يجب أن تعتمد درجات نصفى الاختبار على مجموعات من الأسئلة ذات زمن محدد منفصل بعضها عن بعض • ويمكن تحقيق ذلك باستخدام جزئين متكافئين من الاختبار يعطى كل جزء منهما فى نصف الزمن الكلىللاختبار • وهذه الطريقة أقرب الى طريقة تطبيق الصور المتكافئة من الاختبار فى نفس الجلسة أو بعدها مباشرة ، الا أن كل صورة هى نصف اختبار فى حين أن الدرجة التى يحصل عليها المفحوص هى درجته فى الاختبار الكلى (أى فى نصفيه) ، ولذلك لابد من اســتخدام معادلة سبيرمان - براون أو غيرها من المعادلات المناسبة لايجاد ثبات الاختبار ككل •

واذا استحال تطبيق نصفين للاختبار مستقلين فيمكن تقسيم الزمن الكلى الى أكثر من نصفين ثم اعطاء درجة لكل قسم • ويمكن أن يتم ذلك بأن نطلب من المفحوص أن يضع علامة معينة امام السؤال الذى يكون قد وقف عنده عندما يعطى الممتحن اشارة معينة متفق عليها مقدما • ويمكن أن نجمع عدد المفردات التى اجاب عليها المفحوص اجابة صحيحة فى القسمين الأول

والرابع مثلا وتعتبرها مجموعة مستقلة ثم نجمع اجابات القسمين الثانى والثالث فى مجموعة أخرى بحيث يمثل كل من هاتين المجموعتين نصف اختبار . ومثل هذا النوع من التقسيم الرباعى يحقق نوعا من التوازن بين الآثار التراكمية للممارسة والتعب وغيرها من العوامل . وتصلح هذه الطريقة فى الاختبارات الأخرى التى لا تعتمد على السرعة وخاصة حينما لا تكون الأسئلة مرتبة حسب مستويات الصعوبة .

والسؤال هو : متى يمكن أن نعتبر الاختبار من النوع الموقوت وبالتالى يجب أن نأخذ فى اعتبارنا ما سبق مناقشته ؟ من الواضح أن مجرد استخدام زمن محدد للاختبار لا يعنى أنه اختبار سرعة لأنه اذا انتهى جميع المفحوصين من الاجابة فى الزمن الذى يحدده الاختبار فان السرعة لا تلعب دوراً فى تحديد الدرجة . وحتى اذا أتم جميع المفحوصين ٤٠ سؤالا من ٥٠ فان الفروق الفردية التى ترجع الى السرعة تصبح لا وجود لها رغم انه لا يوجد مفحوص واحد أجاب على جميع أسئلة الاختبار . ولذلك لا يمكن أن نعتبر مجرد محاولة الاختبار كله او عدم محاولته دليلا على السرعة فى مقابل القوة ، وانما لابد من اللجوء الى وسائل أكثر دقة لتحديد المدى الذى يمكن أن نرجع فيه الفروق الفردية فى الأداء الاختبارى الى السرعة ويتم ذلك بتحديد النسبة التى يمكن أن نرجعها الى التباين فى السرعة من التباين الكلى لدرجات الاختبار . ويمكن تحديد هذه النسبة بسهولة بايجاد تباين درجات الأسئلة التى أجاب عليها جميع المفحوصين وقسمة هذا المقـــــدار على تباين الدرجات الكلية للاختبار . ويمكن صياغة ذلك فى صورة المعادلة الآتية :

$$\text{(ع}^{2}\text{ س) تباين السرعة} = \frac{\text{تباين الدرجات التى تمت الاجابة عليها (ع}^{2}\text{ ص)}}{\text{التباين الكلى لدرجات الاختبار (ع}^{2}\text{ ك)}}$$

فاذا حدث واجاب جميع المفحوصين على ٤٠ سؤالا يصبح بسط المعادلة صفرا حيث لا توجد فروق فردية (أو تباين) فى عدد الأسئلة التى تمت الاجابة عليها (ع^{2} ص = صفر) ، وبالتالى يصبح ناتج المعادلة مساويا للصفر ولا يتحقق ذلك بالطبع فى اختبـــارات القـــوة والتى لا تتـــأثر مطلقـــا بالسرعة . ومن ناحية أخرى اذا كان التباين الكلى للاختبار (ع^{2} ك) يرجع

الى فروق فردية فى السرعة فقط ولا يتأثر مطلقا بالقوة فان تباين الدرجات التى قد تكون تمت الاجابة عليها ($ع^٢$ ص) يصبح مساويا للتباين الكلى ويكون ناتج المعادلة مساويا للواحد الصحيح أى يكون الاختبار اختبار سرعة بحت ٠

وهناك طرق احصائية أخرى عديدة وأكثر دقة لتحديد هذه النسبة ولكن مناقشتها تفصيلا يخرج عن نطاق هذا الكتاب ٠

ويمكن أن نضرب مثلا بأثر السرعة على معاملات التناسق الداخلى من بحث قامت به أنستازى ودربك (فى ١١ : ١٥٢ ـ ١٢٣) على اختبارات القدرات العقلية الأولية لأعمار تتراوح بين ١١ سنة و ١٧ سنة ٠ وفى هذا البحث حسبت معاملات الثبات بطريقة التجزئة النصفية أولا ثم عن طريق حساب معاملات ارتباط الدرجات فى نصفين منفصلين تبعا للسرعة والزمن وتتلخص النتائج فى الجدول رقم (٢) :

جدول رقم (٢) معاملات الثبات لبعض اختبارات بطارية القـــدرات العقلية الأوليــة

طريقة حساب معامل الثبــات	معانى الكلمات	الاستدلال	المكان	العدد
التجزئة النصفية فى محاولة واحدة	٠ر٩٤	٠ر٩٦	٠ر٩٠	٠ر٩٢
استخدام نصفين مستقلين من حيث الزمن	٠ر٩٠	٠ر٨٧	٠ر٧٥	٠ر٨٣

ومن هذا الجدول يتضح أن اختبار معانى الكلمات هو فى أساسه اختبار قوة ، بينما يعتمد اختبار الاستدلال على السرعة الى حد ما ، أما اختبار المكان فيعتمد اعتمادا كبيرا على السرعة ٠ كما يتضح أيضا أننا عنـــدما نحسب معاملات الثبات بطريقة جيدة فان هذه المعاملات تنخفض فى اختبار المكان من ٠ر٩٠ الى ٠ر٧٥ ، وكذلك الشأن فى اختبارى الاستدلال والعدد فان معاملات ثباتها تنخفض نسبيا ، أما اختبار معانى الكلمات الذى لا يعتمد على السرعة فان معامل ثباته لم يختلف اختلافا له دلالة عندما حسب بكل من الطريقتين آنفتى الذكر ٠

أما اذا حسب معامل ثبات الاختبارات الموقوتة بطريقة التجزئة النصفية، ثم صحح أثر التجزئة بمعادلة سبيرمان - براون فان جلكسون Gulliksen يقترح تصحيحا اضافيا بالمعادلة الآتية :

ر = رَ - (م ت / ع^٢ خ)

حيث الرمز ر = معامل الثبات بعد تصحيح أثر السرعة

رَ = معامل الثبات بعد تصحيح سبيرمان -- براون

م ت = متوسط الأسئلة المتروكة

ع^٢خ = تباين الخطأ (تباين الاجابات الخاطئة والاجابات المحذوفة)

٢ - مدى الفروق الفردية فى العينة :

تمثل طبيعة الفروق الفردية فى عينة الأفراد التى تطبق عليها الاختبارات بفرض تقدير معاملات الثبات عاملا هاما مؤثرا فى هذه المعاملات . ويجب أن نؤكد أن معامل الارتباط عموما يتأثر بمدى الفروق الفردية فى الأفراد ، فاذا انعدمت هذه الفروق بالنسبة لخاصية من الخصائص فان معامل ارتباط هذه الخاصية بغيرها يصبح صفرا ، ويكون من المستحيل علينا أن نتنبأ بمركز الفرد فى أى خاصية أخرى من معرفة درجته فى الاختبار الذى وضعناه . واذا أخذنا مثالا على ذلك أقل تطرفا وليكن معامل الارتباط بين درجات اختبار للفهم اللغوى واختبار للاستدلال الحسابى طبقا على عينة متجانسة تجانسا شديدا (مثلا طلاب احدى كليات الجامعة) فان معامل الارتباط فى هذه الحالةقد يكون منخفضا انخفاضا واضحا دون أن يدل ذلك على ضعف العلاقة بين الاختبارين . أما اذا أعطينا الاختبارين لعينة غير متجانسة فان معامل الارتباط بينهما قد يكون أعلى منه فى الحالة الأولى .

ويوضح لنا ذلك اثر تحديد مدى الفروق الفردية على معامل الارتباط بوجه عام • ويطلق على هذه الظاهرة فى علم الاحصاء اسم المدى المحدود للعينة restricted range ، ويعتمد ذلك فى جوهره على مدى التباين فى العينة •

وهذا القول يصدق فى عمومه على معاملات الثبات ـ باعتبارها فى الأصل معاملات ارتباط ـ وتوجد معادلات كثيرة لتقدير معامل الثبات المتوقع عندما يزيد أو يقل الانحراف المعيارى فى الجماعة •

٣ ـ خصائص العينة :

يختلف معامل الثبات بين الجماعات التى يتفاوت متوسط مستويات القدرة فيها • وهذه الفروق لا يمكن التنبؤ بها أو تقديرها بأى معادلة احصائية، ولكن يمكن اكتشافها وتحديدها بتجريب الاختبار على جماعات تختلف فى العمر أو فى مستوى القدرة • ومثل هذه الفروق قد تنشأ من أن المستويات المختلفة لصعوبة الاختبار تقيس فى الواقع قدرات مختلفة نوعا ما بعضها عن بعض ، أو قد تنشأ من الخصائص الاحصائية لنوع الدرجة الذى نستخدمه فى تصحيح الاختبار كما هو الحال فى درجة نسبة الذكاء فى اختبار ستانفرد ـ بينه مثلا ، حيث نجد أن معاملات الثبات قد تكون منخفضة نسبيا فى الجماعات الأصغر عمرا أو الأقل قدرة لأن هذه الدرجات تتأثر فى هذه الجماعات بالتخمين ، ولذلك لا يجب تطبيق الاختبار فى هذه المستويات •

ومن الواضح أن كل معامل ثبات يجب أن يصحبه وصف دقيق لنوع المجموعة التى حسب من درجات أفرادها • ويجب أن ننبه على وجه الخصوص الى تباين القدرة ومستوياتها فى العينة • ولا يكون معامل الثبات الذى يتضمنه دليل الاختبار صالحا الا للجماعات المشابهة لعينة التقنين • ولذلك يحسن أن نقسم عينة التقنين هذه الى مجموعات فرعية أكثر تجانسا بالنسبة لعوامل الجنس والعمر والصف الدراسى (أو المرحلة التعليمية) والمهنة وغيرها ، ثم حساب معاملات ثبات منفصلة لكل من هذه المجموعات • وفى هذه الحالة يمكن الاستفادة من معاملات الثبات التى حسبت من عينات التقنين حين نستخدم الاختبار فى الأغراض التطبيقية والعملية •

٤ - عدد أسئلة الاختبار :

يزداد معامل ثبات الاختبار مع زيادة عدد الأسئلة التي تكونه ، ولعل هذا العامل لا يحتاج الى مزيد من التفصيل بعد مناقشتنا المستفيضة السابقة لتصحيح أثر تجزئة الاختبار فى طريقة التجزئة النصفية ، حيث يفترض أن معامل ثبات نصف الاختبار أقل من معامل ثبات الاختبار الكلى •

٥ - تباين الخطـــأ :

يؤثر فى ثبات الاختبار أى مصدر لتباين الخطأ يكون فعالا فى الموقف الاختبارى ، ومفهوم الثبات يعتمد فى جوهره كما بينا آنفا على مدى تحرر الاختبار من هذه المصادر ، وعلى الباحث النفسى أن يجرب كل سبيل ممكن للتأكد من ذلك ولعل طرق حساب الثبات التي عرضناها تهيىء له فرصـا متعددة متنوعة للوصول الى هذا الهدف •

الخطأ المعيارى للقياس :

يمكن التعبير عن مفهوم ثبات الاختبار فى ضوء الخطأ المعيارى للقياس (ع ق) ، أو الخطأ المعيارى للدرجة فى الاختبار • ويفيد هذا المفهوم فائدة كبرى فى تفسير درجات المفحوصين كأفراد ، ولذلك تفوق أهميته فى بعض الحالات مفهوم الثبات نفسه• ويمكن حساب الخطأ المعيارى للقياس من معامل الثبات مباشرة بالمعادلة الآتية :

$$ع_ق = ع_ت \sqrt{١ - ر}$$

حيث ع ق = الخطأ المعيارى للقياس

ع ت = الانحراف المعيارى لدرجات الاختبار

ر = معامل الثبات

بشرط أن يحسب ع ت ، ر لنفس العينة •

ولتوضيح ذلك نعطى مثالا : نفرض أن الانحراف المعيارى لأحـــد

اختبارات القدرة الموسيقية ١٠ ومعامل ثباته ٨٤ر فاننا نحصل على الخطأ المعيارى للقياس مباشرة بالمعادلة السابقة كما يلى :

$$ع_ق = ١٠ \sqrt{١ - ٨٤ر} = ٤$$

ويمثل هذا المقدار الدرجة التى تتذبذب فيها درجات المفحوص اذا افترضنا أنه تعرض للقياس بنفس الاختبار عددا كبيرا من المرات وأن درجاته تتوزع توزيعا اعتداليا ، أى أنه يعبر عن الانحراف المعيارى لدرجات المفحوص الواحد فى نفس الاختبار ، لو أعيد تطبيق هذا الاختبار عليه عشرات المرات ٠

لنفرض أن أحد المفحوصين حصل على الدرجة ٤٥ فى هذا الاختبار ونريد أن نحدد الحدود الدنيا والعليا التى تتوقع لدرجاته الا تتجاوزها لو أعيد عليه القياس (حدود الثبات) ٠ فى هذه الحالة يقوم الباحث بما يلى :

١ ـ اختيار احدى نقاط المنحنى الاعتدالى ولتكن النقطة التى تزيد على المتوسط (م) بما يعادل + ٢ر٥٨ ع ، وهى النقطة التى تحددها نظرية المنحنى الاعتدالى بأن عندها يمكن أن يتم استيعاب ٩٩٪ من درجات هذا المفحوص لو أعيد تطبيق الاختبار عليه عشرات المرات ، أما النسبة المئوية الباقية (وهى ١٪) فتمثل احتمال حدوث خطأ القياس وعدم حصول المفحوص على الدرجة ، أى أن حدوث احتمال حدوث الخطأ فى هذه الحالة ± ١٪ ٠

٢ ـ حساب النقطة التى تقابل (٢ر٥٨ ع) فى درجات المفحوص بافتراض أن الخطأ المعيارى (ع ق) هو فى جوهره انحراف معيارى (ع) وفى مثالنا الحالى تكون هذه النقطة = ± ٢ر٥٨ × ٤ = ± ١١ر٣٢ والتى قد تكون موجبة أو سالبة حسب موضع الانحراف المعيارى من المتوسط٠

ومن هذا المثال يمكن أن نستنتج أن درجة هذا المفحوص عند تطبيق الاختبار عليه مرة واحدة سوف تمتد بين (٤٥ ـ ١١) و (٤٥ + ١١) أى بين ٣٤ ، ٥٦ بدرجة من اليقين يصل الى ٩٩٪ ٠ فاذا طبق عليه هذا الاختبار

أو أى اختبارات مكافئة مائة مرة فان درجته التى يمكن أن تقع خارج هذا النطاق لن تحدث الا مرة واحدة ، ويمثل هذا النطاق حدود الدرجة الحقيقية true score للفرد .

ومن الواضح أن الخطأ المعيارى للقياس يصف الأداء الفردى ويفسره بينما يعبر معامل الثبات عن الأداء الجماعى ، ولهذا لا تصلح الأخطاء المعيارية للمقارنة بين الاختبارات فى درجة الدقة ، وفى هذه الحالة لا تصلح الا معاملات الثبات .

ولعل من أهم فوائد الخطأ المعيارى للقياس - الى جانب تفسيره لأداء الفرد - أنه يجعل الفاحصين لا يبالغون فى تقدير الفروق الضئيلة بين الدرجات . وهذا الحذر يفيد الفاحص سواء عند المقارنة بين درجات مفحوصين مختلفين فى الاختبار الواحد أو المقارنة بين درجات نفس المفحوص فى اختبارات مختلفة . وكذلك يجب فهم الدرجات التى يحصل عليها المفحوصون بعد المعالجات التجريبية فى ضوء الخطأ المعيارى للقياس . الا أننا فى هذه الحالة نحتاج إلى حساب الخطأ المعيارى للفروق بين المقاييس باستخدام المعادلة الآتية :

$$ع_ف = \sqrt{ع^2_{ق_1} + ع^2_{ق_2}}$$

حيث

ع ف = الخطأ المعيارى للفرق بين مقياسين

$ع_{ق_1}$ ، $ع_{ق_2}$ = الخطأ المعيارى للمقياس الأول والمقياس الثانى على التوالى .

ويمكن اعادة صياغة المعادلة السابقة باستخدام معاملات الثبات كما يلى .

$$ع_ف = ع\sqrt{2 - ر_1 - ر_2}$$

مع ملاحظة أن الرمز ع يستخدم للاختبارين معا لأن درجاتهما يجب أن

تحول الى نفس المقياس قبل المقارنة بينهما (أى يتم تحويل الدرجات الخام الى درجات معيارية أو درجات معيارية معدلة على النحو الذى سنبينه فى الفصل السابع) • ويوضح ذلك المثال الآتى :

نفرض أننا طبقنا اختبارين معامل ثباتهما ٩٧ر ، ٩٢ر على التوالى، وأن درجاتهما الأصلية الخام حولت الى نسب ذكاء انحرافية (درجات معيارية معدلة) متوسطها ١٠٠ وانحرافها المعيارى ١٥ ، يمكن أن نحسب الخطأ المعيارى للفرق بين درجتى هذين الاختبارين بالتعويض فى المعادلة السابقة كما يلى :

عف = ١٥ √ ٢ − ٨٧ر − ٨٣ر = ٨ر١

ويمكن أن نستخدم هذه القيمة بنفس المعنى الذى استخدم فيه الخطأ المعيارى للمقياس الواحد •

الفصـــل الخامس

صدق الاختبارات

معنى صدق الاختبار :

يتعلق موضوع صدق الاختبارات Validity بما يقيسه الاختبار والى أى حد ينجح فى قياسه ويتصل هذا بمدى وصولنا الى تنبؤ دقيق أو استنتاج صحيح من الدرجة التى يحصل عليها المفحوص فى الاختبـــار • وهذا الموضوع لا يقتصر على عملية القياس وحدها وانما قد يمتد الى المنهج التجريبى بصفة عامه والذى يتطلب اختبار فروض علمية حول طبيعة الاختبار وما يقيســـه •

واذا كان الصدق مسألة على درجة كبيرة من الأهمية فى تقرير صلاحية أو فائدة أى فرض علمى أو فى تحديد معنى مفهوم معين فانه فى الواقع هو الذى يحدد قيمة الاختبار ومغزاه • ويجب أن نقرر منذ البداية أن عنوان الاختبار لا يدل بالضرورة على ما يقيسه الاختبار لأنه عادة ما يكون مجرد كلمة أو عبارة قصيرة تهدف الى التعرف على الاختبار • ولا يمكن أن نستدل على العملية النفسية التى يقيسها الاختبار الا بفحص مجموعة من المعلومات والبيانات التى تتوافر لنا بالطرق الصحيحة وتكون مســـتقلة نســـبيا عن الاختبار •

طرق تحديد الصدق :

يمكن القول بوجه عام أن جميع طرق تحديد صدق الاختبارات تبحث عن العلاقة التى تربط أداء الأفراد فى الاختبار وبعض الحقائق والمعلومات التى تتصل بالعملية النفسية موضوع الاختبار ، والتى (أى هذه المعلومات والحقائق) يمكن ملاحظتها ملاحظة مستقلة • وهناك كثير من الطرق التى تستخدم فى هذا الصدد ولها مسميات كثيرة ، الا أننا سوف نتحدث عن الأنواع الأربعة من الصدق التى وصفها كتيب التوصيات الفنية للجمعيـــة الأمريكية لعلم النفس (١٤١٣) والتى عالجتها مؤلفات انســـتازى (١١) وكرونباك (٢٠) بالتفصيل •

١ ـ صدق المحتوى :

يقصد بصدق المحتوى Content Validity أن يقوم الباحث النفسي بفحص مضمون الاختبار فحصا دقيقا منتظما لتحديد ما اذا كان يشتمل على عينة ممثلة لميدان السلوك الذى يقيسه كما هو الحال فى الاختبارات المرجعة الى المعيار أو تحتوى على معظم ان لم يكن جميع جوانب هذا السلوك ، كما هو الحال فى الاختبارات المرجعة الى المحك ٠ ومثل هذا النوع من الصدق نستخدمه فى العادة فى تقويم الاختبارات التحصيلية التى نعدها لقياس مقدار ما تمكن التلاميذ من اتقانه من جوانب مهارة معينة أو ما حصلوه من مقرر دراسى أو برنامج مهنى ٠ وقد يبدو أن مجرد فحص مضمون الاختبار يكفى للتدليل على صدقه وصلاحيته للغرض الذى وضع من أجله ، فاختبار معين فى الضرب أو التاريخ قد يبدو صادقا بالتعريف اذا اشتمل على أسئلة تدور حول عملية الضرب أو التاريخ ، الا أن ما نحب ان ننبه اليه ان هذا قد لا يكون صحيحا ، فليس مجرد تأمل محتوى الاختبار يدل على صدقه ٠

ولكن الأمر ليس على هذا القدر من السهولة لأن من أهم المشكلات التى نواجهها هنا هى مشكلة عينة المحتوى أو أصله الاحصائى العام ٠ فنطاق المحتوى الذى نختبره يجب أن يتم تحليله تحليلا دقيقا منتظما حتى نتأكد من أن جميع الجوانب الأساسية تغطيها أسئلة الاختبار تغطية ملائمة وبالنسب الصحيحة حتى لا يكون لدينا عدد كبير من الأسئلة فى جانب معين على حساب الجوانب الأخرى ٠ ويتطلب ذلك بادىء ذى بدء ، وصفا تفصيليا دقيقا لدى المهام والمثيرات والمواقف والمواد التى يستوعبها نطاق المحتوى ،كما يتطلب تفصيل أنواع الاستجابات التى يتوقع لها أن تصدر عن المفحوص ويجب أن يهتم بها الفاحص أو الملاحظ أو المصحح ، والتى قد تدل على عمليات نفسية معينة ، وكذلك توصيف الأهداف التعلمية أو التدريبية أو العلاجية بشرط أن تصاغ فى عبارات اجرائية سلوكية تقبل الملاحظة والقياس ، وبهذا يكون صدق المحتوى أكثر اعتمادا على مدى الارتباط بين استجابات المفحوص على أسئلة الاختبار وميدان السلوك الذى نختبره ، وأى تغيير فى أحد هذه العناصر الثلاثة يغير مباشرة فى طبيعة الاختبار ٠ ففى قياس القدرة على القراءة مثلا قد يلجأ الباحث ـ بالنسبة لأهداف معينة ـ الى بناء اختباره فى حدود المادة القرائية التى تقدمها المدرسة ٠ وقد يلجأ ـ بالنسبة لأهداف

أخرى - الى أن يتناول الاختبار نطاقا أوسع من المواد المقروءة مثل الروايات والصحف والمجلات • ولهذا لابد لمؤلف الاختبار من أن يحــــدد فى دليل اختباره وبدقة عناصر التحليل الثلاثة التى أشرنا اليها (المواد والاستجابات والأهداف) حتى يمكن لأى فاحص أن يحكم على مدى ملاءمة الاختبار لأهدافه فى التقويم النفسى •

وصدق المحتوى يفيد كما سبق أن قلنا فى اختبارات التحصيل ، فهذه الاختبارات يكون لمحتواها - فى العادة - أصل احصائى يمكن أن يعبر عنه كله فى الاختبار ، أو تنتقى منه عينة ممثلة ، وخاصة فى الاختبـــارات التحصيلية التى تستخدم فى تقويم مدى الوصول الى الأهداف التعليمية من المستويات المنخفضة • الا أن صـــدق المحتوى لا يكون كافيا لاختبـارات الاستعدادات ومقاييس الشخصية بل ولاختبارات التحصـــيل التى تقيس العمليات المعرفية العليا ، بل قد يكون مضللا فى كثير من الحالات • فعلى الرغم من أن هذا النوع الأخير من الاختبارات لابد ان يتوفر فيها محكا الملاءمة والفاعلية بالنسبة لمحتواها وخاصة فى المراحل الأولى لاعدادها الا أن هذا وحده لا يكفى ، بل لابد من أن نلجأ بعد ذلك الى وسائل التحقق التجريبى المضبوط ، التى سوف نتحدث عنها فيما بعد ، لتقدير صدقها • ويتضح ذلك اذا علمنا أن اختبارات الاستعدادات والعمليات المعرفية العليا والشخصية ، على عكس الاختبارات التحصيلية ، لا تتشابه تشابها داخليا كبيرا مع ميدان السلوك الذى تعد عينة منه ، وبذلك لا يكون لمحتواها أى دور فى التحقق من الفروض التى يصوغها مؤلف الاختبار عند انتقاء محتواه ، لأن هذه الفروض تتطلب فى هذه الحالة تحقيقا تجريبيا منظما حتى يمكن التأكد من صـــدق الاختبار • وبعبارة أخرى نستخدم الاجراءات التجريبية المنظمة للتحقق من درجة التشابه بين عينة السلوك الذى يتضمنها اختبار الاستعداد والسلوك الذى تنتمى اليه •

ومن ناحية أخرى فان اختبارات الاستعداد والشخصـــية على عكس الاختبارات التحصيلية لا تعتمد على مقرر دراسى أو على برنامج تدريبى أو على مجموعة موحدة من الخبرات السابقة يمكن منها صياغة محتوى محدد للاختبار ، وانما يتفاوت الأفراد كثيرا فى طرق عملهم وفى العمليات النفسية التى يستخدمونها عند الاجابة على نفس الأسئلة • وعلى ذلك فان الاختبار

الواحد قد يقيس عمليات سلوكية مختلفة فى الأفراد المختلفين · وفى مثل هذه الأحوال يستحيل علينا تحديد هذه العمليات التى يقيسها الاختبار من مجرد فحص محتواه · فقد يحل طالب الجامعة مشكلة معينة مستخدما التعبير اللغوى (أو الرمزى عموما ومنه الرياضى) ، بينما يحل الميكانيكى نفس المشكلة ويصل الى نفس الحل مستخدما التصور البصرى المكانى · كما أن اختبارا معينا للاستدلال الحسابى لطلاب المرحلة الاعدادية قد يقيس مجرد الفروق الفردية فى سرعة العمليات العددية عند طلاب الجامعة ·

ويجب ألا نخلط بين صدق المحتوى والصدق الظاهرى face validity فالصدق الظاهرى ليس صدقا بالمعنى العلمى للكلمة لانه يدل على ما يبدو أن الاختبار يقيسه من الظاهر لا على ما يقيسه بالفعل · كما انه يعنى ما اذا كان الاختبار يبدو صادقا فى نظر كل من المفحوصين والفاحصين · والواقع أن موضوع الصدق الظاهرى يدخل فى باب العلاقة الطيبة التى يكونها الفاحص مع المفحوص ، وهو من الأمور المرغوب فيها دون شك رغم ما يسببه من خلط وتداخل مع مفاهيم الصدق الحقيقية فمما لا شك فيه أنه اذا بدا الاختبار للفاحصين أو المفحوصين لا معنى له ، أو لا علاقة له بالظاهره موضوع القياس ، أو غير مناسب أو ساذجا أو سخيفا ، فمن المحتم أن تعاون هؤلاء سيكون ضعيفا حتى ولو كان الصدق الحقيقى للاختبار عاليا · بعبارة أخرى لابد أن يتوافر فى الاختبار هنا الصدق الظاهرى حتى يكون أكثر فاعلية فى مواقف القياس الفعلية ·

ويمكن أن يتحسن الصدق الظاهرى للاختبار بمجرد اعادة صياغة أسئلته بحيث تبدو أكثر ارتباطا بموقف القياس الفعلى · فمثلا اذا كان لدينا اختبار للاستدلال الحسابى البسيط نريد أن نطبقه على مجموعة من المتقدمين للمهن الميكانيكية يحسن أن تستبدل أسئلة الحساب التقليدية التى تتضمنها الكتب المدرسية بأسئلة أخرى أكثر ارتباطا بالعمل الميكانيكى دون أن يحدث أى تغيير يذكر فى الوظيفة السلوكية التى نقيسها · ومع ذلك فان الصدق الظاهرى ذاته لا يمكن أن يكون بديلا لأنواع الصدق الأخرى التى تتحدد بالوسائل الموضوعية ، كما أن تحسين الصدق الظاهرى لا يؤدى بالضرورة الى تحسين الصدق الموضوعى ·

طرق تحديد صدق المحتوى :

توجد عدة طرق لتحديد صدق المحتوى نذكر منها ما يلى :

١ - جـدول المواصـفات : يسـاير هـذا النـوع من الصـدق عمليـة بنـاء الاختبـارات خطـوة خطـوة ، لأن اعـداد اسـئلة أو مفردات هذه الاختبارات انما يتم عن طريق الفحص الدقيق الذى يعتمد على تحليل المحتوى للمثيرات والاستجابات والأهداف (التى فصلناها آنفا) ، واستشارة الخبراء واللجوء الى المحكمين · وعلى أساس المعلومات التى يتم تجميعها يتم تحديد مواصفات الاختبـار Test specifications والتى تتضمن الموضوعات التى يجب أن يشملها الاختبار ، والاستجابات التى يجب اختبارها · والطريقة الأكثر يسرا فى تحديد مثل هذه المواصفات هى اعداد جدول ثنائى يسمى جدول المواصفات table of specifications أو مخطط blue print · ولأن هذا الجدول أكثر شيوعا فى الاختبـارات التحصيلية نعطى مثالا عليه من هذا الميدان الا أنه لا يوجـد ما يمنع من استخدامه فى بناء الاختبارات من الأنواع الأخرى ·

جدول مواصفات لعينة المحتوى والسلوك (الأهداف) : يلعب جـدول المواصفات دورا هاما فى الاختبارات المرجعة الى المعيار والاختبارات المرجعة الى المحك ، والفرق بين هذا الجدول فى الحالتين أنه فى النـوع الثانى من الاختبارات يشتمل على كل الأصل الاحصائى للعناصر الهامة للمحتـوى والسلوك (الأهداف) ، اما فى النوع الأول فان هذا الجدول يمثل عينة من كل منهما ·

ويتم اعداد هذا الجدول فى الحالتين عن طـريق الفحص الدقيق للمقررات والكتب الدراسية من ناحية ، ونواتج التعلم التى يجب اختبارها (فى ضوء الأهداف) من ناحية اخرى ، ثم تحـديد الأهمية النسـبية للموضوعات والأهداف ، وعلى هذا الأساس يتم تحديد عدد المفردات التى ترتبط بكل هدف فى كل موضوع ، وذلك باعداد جـدول ثنائى (جـدول مواصفات) توضع فيه الأهداف التعليمية افقيا والموضوعات رأسيا ، ويعطى الجدول رقم (٣) مثالا على ذلك · ومنه يتضح عدد المفردات التى يشـملها الاختبار بالنسبة لكل موضوع من موضوعات المحتوى فى علاقته بكل هدف.

من الأهداف التعليمية (التى قد تكون أهدافا معرفية أو وجدانية أو مهارية حركية) وقد ملئت خانات الجدول على أساس الوزن النسبى لكل هدف (فى صورة النسب المئوية الموجودة فى السطر الأخير من الجدول) ، والوزن النسبى لكل موضوع (فى صورة النسب المئوية الموجودة فى العمود الأخير من الجدول) ، وهذه الأوزان النسبية يجب تحديدها مقدما ، والطريقة المتاحة حتى الآن لتحديد الوزن النسبى للاهداف هى استطلاع آراء الخبراء (*) ، أما الوزن النسبى للمحتوى فيمكن تحديده موضوعيا فى ضوء الوقت المخصص لتدريسه فعليا ، ويحسن أن تكون فى جميع الحالات فى صورة نسب مئوية ٠

وقد ملئت خانات الجدول رقم (٣) فى صورة تكرارات متوقعة ٠ويقترح فؤاد أبو حطب (٩) المعادلة الآتية لتحقيق هذا الغرض ٠

$$ز = \frac{ا \times ب}{ن}$$

حيث يدل الرمز (ز) على الوزن النسبى للخانة ٠

(أ) على الوزن النسبى للهدف (العمود) الذى تقع فيه الخانة فى صورة نسبة مئوية ٠

(ب) على الوزن النسبى لعنصر المحتوى (السطر) الذى تقع فيه الخانة فى صورة نسبة مئوية ٠

(ن) على المجموع الكلى للاوزان (وهو هنا ١٠٠) ٠

(*) نحب أن ننبه الى ابتذال طريقة استطلاع آراء الخبراء واللجوء الى المحكمين فى عدد كبير من الاختبارات التى ظهرت فى السنوات الاخيرة فى مصر وبعض البلاد العربية ، الى الحد الذى يدفعنا الى الشك فى كثير من بيانات الصدق التى تجمعت منها ، والملاحظ أن كثيرا من الباحثين يلجأون اليها للحصول على معلومات شبه امبريقية من أى نوع وبأى سبيل ربما لسهولتها الظاهرة ويسرها النسبى ، وفى رأينا أن الباحثين فى حاجة الى بذل جهد اكبر فى مجال صدق المحتوى لاختباراتهم ٠

جدول رقم (٢)

جدول الأوزان النسبية لاختبار تحصيل موضوعى لوحدة فى الجغرافيا
بمرحلة التعليم الأساسى

أوزان المحتوى فى صورة نسب مئوية ↓	أوزان الأهداف فى صورة نسب مئوية ←				مجموع أوزان المحتوى ↓
	معرفة		فهم	مهارة فى تفسير الخرائط الجوية	
	مصطلحات ورموز	حقائق نوعية	تاثير كل عامل من العوامل المختلفة فى تغيير الطقس		
الضغط الجوى	٥	٥	٧	٥	٪٢٢
الرياح	٨	٦	١٠	٨	٪٣٢
	٤	٣	٥	٤	٪١٦
الرطوبة والمطر	٥	٤	٦	٥	٪٢٠
السحاب	٢	٢	٤	٢	٪١٠
مجموع أوزان الأهداف	٪٢٤	٪٢٠	٪٣٢	٪٢٤	٪١٠٠

وقد حصلنا على القيم الموجودة فى خانات الجدول رقم (٣) من تطبيق هذه المعادلة ، فمثلاً فى الخانة الأولى من اليمين (معرفة رموز ومصطلحات الضغط الجوى) حصلنا على الوزن النسبى ٥٪ من ناتج $\frac{٢٤ \times ٢٢}{١٠٠}$

والخانة التى تليها فى نفس السطر (معرفة الحقائق النوعية للضغط الجوى) من $\frac{٢٠ \times ٢٢}{١٠٠}$ وهى تساوى ٤٠ر٤ قربناها الى ٥٪ • ويتطلب الامر بالطبع بعض التقريب حتى نحصل على المجموع الكلى للاوزان وهو عادة ما يكون ١٠٠ ، اذا كانت احكام الخبراء على الأوزان النسبية للاهداف وعناصر المحتوى فى صورة أجزاء من ١٠٠ ٪ •

ويدل الوزن النسبى لكل خانة من خانات الجدول السابق على النسبة المئوية لعدد المفردات التى يجب أن يشملها الاختبار من النوع الذى تدل عليه الخانة • فمثلا اذا كان اختبارنا يتألف من ٥٠ سؤالا فيجب أن يكون ٥٪ من هذه الأسئلة من النوع الذى يقيس معرفة رموز ومصطلحات الضغط الجوى، ٦٪ من هذه الأسئلة تقيس فهم العوامل المؤثرة فى الرطوبة والمطر ، ٨٪ منها يقيس المهارة فى تفسير خرائط الرياح ، وهكذا •

ويوضح الجدول رقم (٤) عدد الأسئلة المتوقع من كل نوع فى كل خانة على اساس أن يكون العدد الكلى لأسئلة الاختبار ٥٠ سؤالا ، واعتمادا على الأوزان النسبية التى حصلنا عليها فى جدول (٣) • فمثلا نتوقع أن يتضمن الاختبار ٣ أسئلة من النوع الذى يقيس معرفة رموز مصطلحات الضغط الجوى على اساس أن هذا العدد هو ٥٪ من المجموع الكلى لأسئلة الاختبار أى $\frac{٥٠ \times ٥}{١٠٠}$ = ٥ر٢ قربناها الى ٣ ، وهكذا •

جدول رقم (٤)

جدول مواصفات اختبار تحصيلى موضوعى لوحدة فى مادة الجغرافيا بالمرحلة الاعدادية يتألف من ٥٠ سؤالا على أساس الأوزان النسبية فى الجدول رقم (٣)

المحتوى	معرفة		فهم تأثير كل عامل في تغيير الطقس	المهارة فى تفسير خرائط الطقس	المجموع الكلى للاسئلة
	رموز ومصطلحات	حقائق نوعية			
الضغط الجوى	٢	٢	٤	٣	١٣
الرياح	٤	٣	٥	٤	١٦
الحرارة	٣	١	٢	٢	٧
الرطوبة والمطر	٢	٢	٣	٢	٩
السحاب	١	١	٢	١	٥
المجموع الكلى للاسئلة	١٢	١٠	١٦	١٢	٥٠

ويفيد جدول المواصفات فى بناء الاختبار وتحديد صدق محتواه من خلال ما يلى :

(أ) تحديد عدد المفردات أو الأسئلة التى يتألف منها الاختبار وقد عرضنا طرق تحديد ذلك فيما سبق ٠

(ب) تحديد نوع المفردات ، فالاختبار من متعدد ليس هو النمط الوحيد مثلا الذى يصلح لقياس جميع الأهداف وجميع جوانب موضوعات المحتوى ، ويمكن للقارىء الرجوع الى الباب الرابع من هـــذا الكتاب ، والمخصص للاختبارات التحصيلية ليتعرف على طبيعة كل نوع من انواع المفردات وما يقيس من عمليات معرفية ، مع التنبيه خاصة الى ما ينشا عن بعض هــذه الأنواع (وخاصة أسئلة البديلين) من أساليب استجابة response set

(جـ) تحديد قالب المهام task format التى يتألف منها الاختبار ٠ ومن المؤكد أن الاختبارات المتشابهة فى القالب (الفاظ أو صور او أعمـال عيانية محسوسة) والمختلفة فى المحتوى (ميكانيكا أو الكترونيات مثلا) تكون فيما بينها أكثر ارتباطا من تلك التى تكون متشابهة فى المحتــوى (ميكانيكا مثلا) وتعرض فى قوالب مختلفة (الفاظ أو صـــور او أعمال عيانية) ، ويقل الارتباط كثيرا بين الاختبارات اذا أختلفت بالطبع فى كل من القالب والمحتوى (٢١ : ١٤٦) ٠ وتبدو أهمية هذه المسألة اذا علمنا أن قياس مهارة حركية مثلا باختبار لفظى يكون أقل صدقا من قياسها باختبار عملى ٠

٢ ـ **بعض الطرق الامبريقية لتحديد صدق المحتوى** : وتوجد عدة طرق امبريقية أخرى للتحقق من صدق المحتوى نذكر منها :

(أ) المقارنة بين الدرجات التى يحصل عليها المفحوصون فى الاختبار، أو فى صورة مكافئة له ، قبل التدريب أو العلاج وبعده للتحقق من التحسن الذى يطرأ على الاختبار نتيجة للتدريب ، ممـــا يؤكد تعلق relevance ما يقيسه الاختبار ببرنامج التدريب أو التعليم أو العلاج النفسى ٠

(ب) دراسة انماط الاخطاء الشائعة فى الاختبار للتأكد من صدق مفتاح الاختبار ٠

(هـ) تحليل طرق العمل التى يستخدمها المفحوصون وذلك باعطاء الاختبار فرديا مع توجيههم الى « التفكير بصوت عال » اثناء حل مشكلات الاختبار ٠

(د) دراسة أثر بعض العوامل غير المرتبطة بمحتوى الاختبار كالسرعة أو القراءة فى علاقتها بما يقيسه الاختبار ، فالسرعة التى لا ترتبط مثلا باختبار يقيس المهارة فى الكتابة على الآلة الكاتبة او اليسر العددى لا ترتبط باختبار آخر يقيس المحصول اللفظى عند الطفل أو قدرته على الاستدلال وحل المشكلات ٠ والقراءة لا ترتبط بقياس المهارة فى التصويب ، كما أن قراءة فقرات مطولة لا ترتبط باختبار يقيس القدرة الكتابية (عند الموظفين الاداريين) وانما تكفى فى هذه الحالة قراءة جمل من المستوى الذى يستخدم فى الحياة اليومية ٠

(هـ) تحديد الخصائص اللفظية والكمية لأسئلة الاختبار أو مفرداته ، وهذه العملية هى جزء من عملية تحليل المفردات item analysis اثناء بناء الاختبار ، وتفيد فى نفس الوقت ـ اذا تمت بنجاح ـ فى تحديد صدق محتوى الاختبار ٠

ويشمل التحليل الكيفى للاسئلة الحكم عليها فى ضوء فعاليتها من حيث صياغتها وصلاحيتها ، وفى هذا الصدد قد يلجأ الباحث الى مجموعة من المحكمين لابداء الراى فى مختلف الجوانب الكيفية للسؤال ٠ أما التحليل الكمى للمفردات فيتضمن على وجه الخصوص تحديد سهولة (أو شيوع) المفردات ، وتمييزها وتحديد صدقها وثباتها ، وهذه جميعا تفيد فى الحكم على مدى ملاءمة محتوى الاختبار ، أو صدق محتواه ٠ وتتناول هذه الجوانب الكمية هنا بشىء من الايجاز ، وعلى القارىء المهتم بالتفاصيل الرجوع الى فصول (تحليل المفردات) فى كتب القياس النفسى أو الاحصاء النفسى المتخصصة ٠

معاملات سهولة (أو شيوع) المفردات : يحسب معامل سهولة المفردة بصفة عامة بالمعادلة الآتية :

$$\text{س} = \frac{\text{س}}{\text{ص} + \text{خ}}$$

حيث س : معامل سهولة المفردة
ص : عدد الاجابات الصحيحة
خ : عدد الاجابات الخاطئة

ويرى فؤاد أبو حطب (١٩ أ) أن هذه المعادلة يمـــكن ان تكون اكثر عمومية حين تستخدم لحساب معاملات شيوع المفردات ·

$$\text{ش} = \frac{\text{أ}}{\text{أ} + \text{ب}}$$

حيث ش = معامل شيوع المفردة
أ = عدد الأفراد الذين أجابوا على المفردة فى الاتجاه (أ) ولیكن نعم فى استخبارات الشخصية أو صـــواب فى اختبارات القدرات أو ذكروا المفردة فى اختبارات الانتاج ·

ب = عدد الافراد الذين اجابوا على المفردة فى الاتجاه (ب) وهو عكس الاتجاه (أ) ·

وقد استخدم فؤاد أبو حطب مفهوم « معامل الشـيوع » Commeness Coefficient of لأول مرة فى منتصف السبعينات فى تقنين اختبار رسم الرجل ، وفى استفتاء للمشكلات الدراسية عند طلاب المرحلة الثانوية ·

دقة مفتاح الاختبار وثبات مفرداته : تصلح لحساب ثبات المفردات أو

الأسئلة الطرق المعتادة لحساب الاختبار ككل ومنها اعادة الاختبار الا أن من افضل الطرق هنا ما يسمى طريقة الاحتمال المتوالى وهى من اقتراح جتمان ومعادلتها هى :

$$ر = \frac{ن - ١}{ن} \left(ل - \frac{١}{ن} \right)$$

حيث ر = معامل ثبات السؤال ·

ن = عدد الاحتمالات الاختبارية للمفردة أو السؤال ، ففى أسئلة البديلين ن = ٢ وفى سؤال الاختيار من بين ٤ بدائل تكون ٤ وهكذا ·

ل = الاحتمال المتولى ، أى أكبر تكرار نسبى لأى احتمال اختيارى من الاحتمالات التى تحتوى عليها المفردة أو السؤال

وهذه الطريقة لا ترتبط بالثبات بمعناه المباشر وهى اقرب الى التحقق من دقة مفتاح الاختبار ·

تمييز المفردات : يمكن استخدام الفرق بين النسبتين المئويتين للافراد الذين اجابوا على المفردة من المستويين الأعلى والأدنى كمؤشر على تمييز

المفردة ويرمز له بالرمز (ف) أو (D).

وتوجد عدة طرق أخرى الاكثر تعقيدا تستخدم مفاهيم المنحنى الاعتدالى لا يتسع المقام لتناولها بالتفصيل ·

الا ان ما يجب أن ننبه اليه أن عددا كبيرا من الاختبارات النفسية لا يتطلب شرط التمييز هذا ، وتظل هذه الاختبارات مع ذلك صادقة من حيث محتواها ، ولعل أشهر الأمثلة المعاصرة على ذلك الاختبارات المرجعية المحك ·

صدق المفردات أو الأسئلة : لعل اهم بيانات تحليل المفردات التى تتصل

مباشرة بصدق محتوى الاختبار ككل ما يسمى صدق المفردات ، والتى تشمل أيضا صدق الاختبارات الفرعية Sub-tests التى يتالف منها الاختبار الكلى ٠

وتوجد طريقتان لتحديد صدق المفردات أو الأسـئلة أو الاختبارات الفرعية همـا :

١ ـ حساب معامل الارتباط بين درجات المفردة أو السؤال أو الاختبار الفرعى ودرجات الاختبار الكلى (والتى تعد فى حالة المحك الداخلى للصدق) أو درجات أى محك خارجى ، ويتحدد نوع معامل الارتباط المحسوب تبعا لطبيعة درجات الجزء (المفردة أو السؤال أو الاختبار الفرعى) من ناحية ودرجات الكل (المحك الداخلى أو المحك الخارجى) من ناحية أخرى ٠

فاذا كانت درجات الجزء من نوع الكم المنفصل الثنائى اى من نوع الكل أو لا شىء (مثل نعم ـ لا ، صفر ـ ١ ، الخ) ودرجات المحك الداخلى أو الخارجى من نفس النوع (مثل نجاح ـ فشل ، تفوق ـ تخلف) يحسب معامل الارتباط الرباعى فى هذه الحالة ٠

أما اذا كانت درجات أحد مكونى العلاقة (الجزء أو المحك) من نوع الكم المنفصل أو الثنائى ، ودرجات المكون الآخر من نوع الكم المتصل يحسب فى هذه الحالة معامل الارتباط الثنائى ٠

أما اذا كانت درجات كل من الجزء (السؤال أو المفردة أو الاختبار الفرعى) والمحك من نوع الكم المتصل يحسب معامل الارتباط التتابعى لكارل بيرسون ٠

٢ ـ طريقة المقارنة الطرفية : وذلك بتقسيم درجات المحك الداخلى أو الخارجى الى مستويين ، ويستخدم هنا أسلوب أشبه بطريقة المجموعات المتضادة تتمثل فى انتقاء مجموعتين متطرفتين من المفحوصين فى ضـوء درجاتهم الكلية فى الاختبار (المحك الداخلى) أو فى ضوء درجا... فى المحك الخارجى ، ويمكن أن تكون هاتان المجموعتان المتطرفتان أولئك ا... ينتسبون

الى الربيع الأعلى (المستوى المرتفع) والى الربيع الادنى (المستوى المنخفض) بالنسبة للدرجات الكلية فى المحك ، ويرى فلاناجان أن انسب مجموعتين متطرفتين يمكن اختيارهما فى ضوء المحك الداخلى خاصة (أى الدرجة فى الاختبار الكلى) هما أولئك الذين يقعون ضمن نسبة ٢٧٪ من المفحوصين الحاصلين على الدرجات العليا فيه ، والذين يقعون ضمن نسبة ٢٧٪ الحاصلين على الدرجات الدنيا ٠

وبعد تحديد المجموعتين المتضادتين فى ضوء المحك الداخلى أو الخارجى تتم المقارنة بين أداء المفحوصين فيهما على المفردة أو السؤال أو الاختبار الفرعى ؛ ويستخدم فى هذا الغرض احدى طريقتين ٠٠

١ ـ فى حالة استخدام نسبة ٢٧٪ العليا ، ٢٧٪ الدنيا يمكن الاعتماد على معاملات الإرتباط التى حسبها فلانا جان فى جداوله الشهيرة والتى تعبر عن العلاقة بين معاملات سهولة السؤال أو المفردة للمجموعتين العليا والدنيا من ناحية ودرجات المفحوصين فى كل من المجموعتين فى السؤال (وهى من نوع الكم المتصل أو الثنائى) وبعبارة أخرى تعبر عن العلاقة بين الدرجة الكلية فى الاختبار ودرجة كل سؤال على حدة من أسئلة الاختبار فى المحك الداخلى أو الخارجى ٠

٢ ـ فى حالة اختيار المجموعتين المتضادتين فى المحك الداخلى أو الخارجى فأى طريقة أخرى غير الطريقة السابقة (كطريقة الربيع الأعلى والأدنى ، أو نسبة ١٠٪ العليا ، ١٠٪ الدنيا مثلا) يمكن حساب دلالة الفروق بين أداء المجموعتين فى المفردة أو السؤال أو الاختبار الفرعى باستخدام احد مقاييس الدلالة الإحصائية المناسب كا٢ أو اختبار «ت» أو تحليل التباين أو غير ذلك من الطرق البارامترية واللابارامترية) ٠

ونحب أن ننبه هنا الى خطأ شاع فى السنوات الأخيرة فى بعض رسائل الماجستير والدكتوراه ، بل وفى بعض « البحوث » التى يتقدم بها أصحابها الى اللجان العلمية الدائمة للترقية الى وظائف أساتذة وأساتذة مساعدين ، ويتلخص هذا الخطأ فى أن الباحث يختار مجموعتيه المتضادتين فى ضوء المحك ثم يقارن بينهما مستخدما درجات المحك ذاتها ، وهذا التحليل قد يفيد فى

تدعيم فكرة « تضاد » المجموعتين فحسب ، ولا يتجاوز الأمر الى أبعد من ذلك، ولا يفيد على أى نحو - بالطبع - فى تقديم اى معلومات عن صدق الاختبار. فالصدق يتطلب - ونكرر - ان يتم تحليل اداء المجموعتين المتضادتين فى المحك على المفردة أو السؤال أو الاختبار الفرعى ·

ويمتد أصحاب الخطأ السابق احيانا الى ما هو اخطر ، فنجدهم يكونون المجموعتين المتضادتين فى ضوء أداء الأفراد فى الاختبار الفرعى ، وليس فى المحك ، ثم يقارنون بين المجموعتين فى درجاتهما فى الاختبار الفرعى نفسه ، باستخدام مقاييس الدلالة الاحصائية ، وبالطبع يحصلون على فروق دالة يستنتجون منها « صدق » الاختبار الفرعى · أليس هذا هو العبث العلمى بعينه ، هو نوع من تحصيل الحاصل ، الذى لا يضيف جديدا الى صدق الاختبار ، ولا يتجاوز حدود « اللعب الاحصائى » السخيف ·

(ب) الصدق المرتبط بالمحكات

يدل الصدق المرتبط بالمحكات criterion-related validity على قدرة الاختبار على التنبؤ بسلوك المفحوص فى مواقف محددة أو تشخيص هذا السلوك · ولهذا لابد من الحكم على الأداء فى الاختبارات فى ضوء أحد المحكات · ويقصد بالمحك مقياس مباشر ومستقل لما يهدف الاختبار الى قياسه والتنبؤ به أو تشخيصه ، أو هو ميزان لتحديد مدى صلاحية الاختبار ، أو هو اختبار للاختبار · وعلى ذلك فبالنسبة لاختبار يقيس الاستعداد الميكانيكى يمكن أن يكون المحك اداء المفحوصين فى عمل ميكانيكى فعلى فى مدرسة ثانوية صناعية أو مصنع · وهنا يجب أن نميز بين غرضين يستخدم فيهما هذا النوع من الصدق : أولهما ، التنبؤ الطويل المدى ، وثانيهما التشخيص ، وذلك فى ضوء العلاقات الزمنية بين المحك والاختبار ، فاذا تصاحبت أو تلازمت زمنيا بيانات المحك ودرجات الاختبار يصبح الصدق فى هذه الحالة من النوع « التلازمى » ، أما اذا وجد فاصل زمنى طويل (قد يصل لعدة سنوات) بين درجات الاختبار ومعلومات المحك فان الصدق يصبح « تنبؤيا » · وهكذا يمكن أن نميز بين نوعى الصدق هذين فى ضوء أهداف القياس · فالمعلومات التى يوفرها الصدق التنبؤى ترتبط بالاختبارات التى تستخدم فى انتقاء الأفراد

وتصنيفهم وتوجيههم تربويا أو مهنيا أو عسكريا وكذلك فى أغراض التنبؤ الكلينيكى · وفى جميع هذه الأغراض تتشابه اجراءات الصدق ، حيث نعطى الاختبار لعينة من الأفراد تمثل المجتمع الأصلى (موضوع الاهتمام) ، ولكن الدرجات التى يحصل عليها أفراد هذه العينة لا تستخدم فى اتخاذ أية قرارات خاصة بهم · كما يجب ألا تعطى هذه الدرجات الى أى شخص أو سلطة يمكن له أو لها أن تتخذ قرارات تؤثر فيهم ، حتى لا يطرأ على المحك أى تشـــويه أو خلط أو تلوث criterion contamination لأن الدرجة الاختبارية فى هذه الحالة قد تؤثر على الوضع النسبى للمفحوص فى المحك ، وخاصة اذا كان من نوع التقديرات · وبعد ذلك تتم متابعة العينة الأصلية حتى تتوافر لدينا عنها المعلومات التى نستقيها من المحك والتى تؤدى بنا الى تحديد المدى الـــذى تتفق فيه درجات الاختبار مع النتائج التى نلاحظها كما تتمثل فى درجات المحك ، وهذا الاتفاق يحدد لنا مدى تنبؤية الاختبار · وتستخدم لهذا الغرض أساليب احصائية معقدة تعتمد فى جوهرها على أسلوب تحليل الانحدار · · regression analysis (راجع ١٩)

وعادة ما يتضمن دليل الاختبار قيما خاصة بالصدق التنبؤى مشتقة من استخدام محكات مختلفة بهدف مساعدة من يستخدم الاختبار فى فهم ما يقيسه وذلك على الرغم من أن مستخدمى الاختبارات قد لا يهتمون اهتماما مباشرا بالتنبؤ بأى من المحكات المستخدمة ·

أما الصدق التلازمى فهو أكثر ملاءمة للاختبارات التى تستخدم لأغراض التشخيص (*) لا التنبؤ بنتائج المستقبل · وفى رأينا أن الغرض الأخير هـــو أكثر الأغراض مشروعية فى اطار المرحلة الراهنة من تطور علم القيـــاس النفسى ·

ويلجأ الباحثون الى الصدق التلازمى فى بعض الأحيان كنوع من حلول الوسط لمشكلة الصدق التنبؤى وما تتطلبه من وقت طويل ، وفيهـــا تعطى الاختبارات لمجموعة من الأفراد تتوافر لدى الباحث عنهم معلومات خاصة

(*) يستخدم فؤاد أبو حطب (٩) مصطلح التشخيص بمعنى أكثر شمولا من معناه الشائع فى اطار علم النفس الكلينيكى ، ويقصد به هنا وصف الوضع الراهن للسلوك وقد يمتد الى الاحداث الماضية ·

بمحك معين ، فنقارن مثلا درجات طلاب الجامعة فى اختبار معين بمتوسط تراكمى لدرجاتهم الدراسية حتى وقت الاختبار كما نقارن درجات العمال فى الاختبار بمستوى نجاحهم الراهن فى العمل •

وتتمثل أهم مشكلات الصدق التلازمى فى صعوبة تحديد علاقات السبب والأثر ، لأنه يدل بالفعل على علاقة اقتران أكثر من دلالته على علاقة «علية» ولذلك فان استخدامه كبديل للصدق التنبؤى يتضمن صعوبات منطقية ظاهرة ، لأن التنبؤ يتضمن فى جوهره هذه العلاقة السببية ، وقد يكون فى انقضاء الفاصل الزمنى بين الاختبار والمحك ما يبرر « العلية » فى الصدق التنبؤى •

الا أن هذه الفروق بين نوعى الصدق ربما يتجاوزها ، ما أدخل على مفهوم معامل الارتباط من توسيع فى معناه وفى وظائفه ، ويدل على ذلك الاستخدامات المتنوعة لأسلوب « تحليل المسار » path analysis (راجع ١٩) ، فالعلية فى هذه الحالة قد تستنتج من البيانات الارتباطية بين الاختبارات والمحكات ، سواء كانت هذه المحكات تنتمى الى الماضى أو الحاضر أو المستقبل •

ومن ناحية أخرى فان للصدق التلازمى استخداماته الخاصة فى أغراض معينة فهو أكثر ملاءمة كما ذكرنا للاختبارات التشخيصية والتى تلجأ اليها فى حل المشكلات الراهنة أكثر من مشكلات المستقبل • فعند اعداد مقياس للشخصية يهتم الصدق التلازمى بسؤال مثل : هل هذا الطفل عصابى ؟ ، أما الصدق التنبؤى فيجيب على سؤال من نوع آخر وهو : هل من المحتمل أن يصير هذا الطفل عصابيا ؟ ومع ذلك فكثيرا ما تثار التساؤلات حول وظيفة الاختبار فى حالة الصدق التلازمى الذى « تتصاحب » فيه معلومات الاختبار ومعلومات المحك ، فالواقع أن الاختبار فى مثل هذه الحالة عادة ما يكون بديلا أبسط وأسرع وأقل تكلفة وجهدا من معلومات المحك • فاذا كان المحك يتطلب ملاحظة مباشرة لمريض مدة أسبوعين من لحظة دخوله المستشفى ، فان الاختبار الصادق صدقا تلازميا يمكنه أن يميز الأسوياء والعصابيين والحالات المشكوك فيها فى وقت قصير ، وبذلك يختصر لنا عدد الأشخاص الذين يتطلبون هذه الملاحظة الشاملة ، بالاضافة الى اقتصاد الوقت والجهد اللذين يبذلان فى الملاحظة المباشرة •

أنواع المحكات : يمكن للاختبار أن يحدد صدقه في ضوء أي عدد من المحكات بعدد الاستخدامات الخاصة لهذا الاختبار · وأي طريقة ملائمة للتقويم النفسي في أي موقف يمكن أن تعد محكا لغرض معين ، بشرط أن تتوافر فيه شروط الاختبار الجيد التي نتناولها في هذا الكتاب · لقد أشرنا فيما سبق الى أن المحك هو ميزان للاختبار أو اختبار للاختبار ، ولهذا لابد من التحوط والحذر في انتقاء المحكات حتى لا توزن الاختبارات بموازين فاسدة أو تختبر باختبارات مضللة · ونعرض فيما يلي لأنواع المحكات الشائعة ·

١ ـ التحصيل المدرسي الأكاديمي العام : هو أشهر المحكات وخاصة بالنسبة لاختبارات الذكاء ، ولذلك توصف هذه الاختبارات عادة بأنها مقاييس الاستعداد المدرسي · ومن مؤشرات هذا التحصيل المدرسي مجموع درجات الاختبارات التحصيلية ، ونتائج النقل وتقديرات التخرج ، ومراتب الشرف ، والجوائز أو المنح ، وتقديرات المعلمين لتحصيل تلاميذهم ، ويستخدم هذا المحك ـ الى جانب اختبارات الذكاء كما أشرنا ـ في حساب صدق بطاريات الاستعدادات الفارقة واختبارات القدرات المتخصصة ، وبعض مقاييس الشخصية ، سواء كان غرضنا من استخدام هذه الاختبارات هو التنبؤ أو التشخيص · وهو محك شائع في جميع المراحل التعليمية وفي جميع أنواع التعليم ·

٢ ـ الأداء في برنامج معين : هذا المحك أكثر ارتباطا بحساب صدق بطاريات الاستعدادات الفارقة واختبارات القدرات المتخصصة ، ومقاييس بعض سمات الشخصية ، وهو يعتمد على مؤشرات أداء المفحوصين في برنامج معين من التدريب المتخصص · ففي حساب صدق اختبار معين للقدرة الميكانيكية يمكن الاعتماد على التحصيل النهائي في مقرر معين يتعلق باحدى المهن الصناعية · وكما يمكن أن نعتبر التحصيل في مقررات التعليم التجاري ، مثل الاختزال والآلة الكاتبة ومسك الدفاتر ، محكات لاختبارات الاستعدادات في هذه الميادين · وكذلك فان أداء طلاب المعاهد الموسيقية والفنية قد يعتبر محكا لحساب صدق اختبارات الاستعدادات الموسيقية والفنية · وقد حسب صدق معظم اختبارات الاستعدادات المهنية باستخدام محكات التحصيل في كليات الحقوق والطب والصيدلة والهندسة وغيرها · أما اذا كانت الاختبارات من النوع الذي نستخدمه في أغراض برنامج معين للانتقاء ، فقد

تستخدم سجلات التدريب كمصدر للمعلومات التى نحتاج اليها ، كما هو الحال فى حساب صدق اختبارات انتقاء الطيارين ، فالمحك هنا هو أداء الطيارين فى مواقف التدريب الفعلية ٠

ومن المؤشرات الخاصة التى يتضمنها محك الأداء فى برامج التدريب يمكن أن نذكر الاختبارات التحصيلية المتخصصة التى نطبقها عقب الانتهاء من برنامج التدريب أو تقديرات المدربين أو النجاح والفشل فى التدريب (الاستمرار والاكمال فى مقابل الانقطاع والاستبعاد) ٠ وقد نستخدم درجات الامتحانات فى مقررات بذاتها فى المدرسة الثانوية أو الكليات الجامعية والمعاهد العليا وذلك لتحديد صدق بطاريات الاستعدادات ، فمثلا يمكن مقارنة درجات الطلاب فى اختبارات الفهم اللغوى بدرجاتهم فى مقررات اللغة العربية ، ودرجات التصور البصرى المكانى بدرجات مقرر الهندسة ، الخ ٠

ولابد أن نميز فى هذا الصدد بين ما يسمى المحكات المتوسطة والمحكات القصوى أو النهائية (١٦) بالنسبة لمؤشرات سجلات التدريب ٠ فعند اعداد اختبار لقياس الاستعداد للطيران أو لقياس الاستعداد الطبى أو الاستعداد للتدريس مثلا نجد أن المحكات النهائية هى الأداء الفعلى فى ميدان القتال ، أو النجاح فى ممارسة الطب ، أو النجاح فى الممارسة الفعلية للتدريس داخل حجرة الدراسة ، ويتطلب ذلك دون شك وقتا طويلا حتى يمكننا الحصول على معلومات من هذه المحكات ، بل انه من المشكوك فيه امكان وصولنا الى محك نهائى فى الواقع ٠ بالاضافة الى انه ، حتى اذا توفر لنا مثل هذا المحك النهائى ، فانه قد يتعرض لعوامل كثيرة يصعب التحكم فيها مما يجعل المحك فى النهاية لا فائدة منه ٠ فمن الصعب مثلا أن تقوم درجة النجاح النسبى للاطباء الذين يمارسون مختلف التخصصات الطبية فى أجزاء مختلفة من الوطن سواء داخل المستشفيات أو عياداتهم الخاصة ٠ ولهذا السبب نلجأ فى العادة الى المحكات المتوسطة التى ذكرنا أمثلة لها فيما سبق مثل سجلات الأداء فى مرحلة معينة من مراحل التدريب ٠

٣ _ الأداء على العمل نفسه : هو أكثر المحكات صلاحية ، لأنه يعتمد على سجلات تتبعية للاداء الواقعى والفعلى للافراد فى العمل ٠ وقد استخدم العلماء هذا المحك الى حد ما فى حساب صدق اختبارات الذكاء العام

والشخصية ، والى حد كبير فى حساب صدق اختبارات الاستعدادات الفارقة والقدرات المتخصصة • ومن المؤكد أن « الأعمال » تختلف فى مستوياتها وفى أنواعها • ويصدق هذا على الأعمال الصناعية والتجارية والمهنية والعسكرية وغيرها ، ورغم أن معظم مقاييس أداء « العمل » قد لا تعتبر محكات نهائية الا انها افضل بكثير من محك سجلات التدريب الذى تحثنا عنه آنفا • ومع ذلك فقد لا يسمح قياس أداء العمل بتوحيد وتقنين الشروط التى تتوافر أثناء التدريب ، كما أنه يتطلب تتبعا طويل المدى مما يؤدى الى قلة عدد المفحوصين، وبالتالى يؤثر فى النتائج على نحو لا حيلة للباحث فيه •

والمؤشرات التى نستخدمها فى قياس درجة النجاح فى « العمل » عديدة منها مقدار الانتاج ، ونوعه ، والاستهداف للحوادث ، والمرتبات والعلاوات، والترقيات والاستمرار فى العمل وعدم الانتقال الى غيره ، وطول مدة الخدمة، ومعدل النمو فى المهنة، وتقديرات المشرفين والرؤساء للكفاءةالمهنية للمفحوص وقد يعتمد مقياس المحك على ملاحظة عينة صغيرة محددة من أداء الشخص فى العمل ، أو قد يشتق من سجل تتبعى للانتاج ، أو تاريخ الشخص فى العمل • بحيث يصبح محكا شاملا للفترة التى يقضيها المفحوص فى عمله •

٤ ـ **المجموعات المتضادة** : محك « المجموعات المتضادة » Contrasted groups ' هو من نوع المحكات المركبة التى تدل على الآثار المتراكمة وغير المضبوطة للحياة اليومية ، أى انه اقل وضوحا وأكثر تعقيدا من المحكات الأخرى التى سبق مناقشتها ، ويعتمد فى أساسه على بقاء الفرد واستمراره داخل مجموعة معينة لفترة طويلة نسبيا من الزمن أو عدم بقائه فيها وخروجه منها • ففى حسابنا لصدق اختبار الذكاء يمكن أن نقارن درجات ضعاف العقول المقيمين باحدى المؤسسات بالدرجات التى يحصل عليها تلاميذ مدرسة عادية فى نفس السن تقريبا • وفى مثل هذه الحالة يتأثر المحك بالعوامل العديدة التى تتحكم فى هذه المؤسسات بوجه عام وتجعلها بيئة مختلفةعن المدرسة العادية • وكذلك اذا أردنا حساب صدق اختبار للاستعداد الموسيقى أو الاستعداد الميكانيكى فيمكن أن نقارن بين الدرجات التى يحصل عليها الطلاب الذين يدرسون الموسيقى أو الهندسة فى كليتين أو معهدين متخصصين ، وبين درجات طلاب آخرين يدرسون فى كلية جامعية لا تعد طلابها اعدادا منظما فى هذين المجالين •

ويمكن للباحث الذى يلجأ الى هذه الطريقة فى حساب الصدق أن يكون المجموعات المتضادة مستخدما ما يشاء من المحكات ، مثل الدرجات المدرسية أو تقديرات النجاح والفشل أو الأداء فى العمل او غيرها ، وذلك باختيار الحالات المتطرفة فى توزيع درجات المحك • وتصبح المجموعات المتضادة فى هذه الحالة هى جماعات تمايزت تدريجيا بفعل مطالب الحياة اليومية وخبراتها العديدة •

وتستخدم طريقة المجموعات المتضادة فى حساب صدق اختبارات الشخصية على وجه الخصوص ، فمثلا يمكن أن نقارن أداء من يمارسون أعمالا قيادية أو اقناعية بأداء من يمارسون أعمالا كتابية فى اختبار معين للسمات الاجتماعية ليمكن حساب صدق هذا الاختبار ، مفترضين أن الأفراد الذين يعملون بمهن معينة تتطلب التعامل الاجتماعى (البيع ـ التدريس •• الخ) وظلوا يقومون بها لفترة من الزمن سوف يتقدمون على غيرهم من الأفراد الذين يعملون فى ميادين أخرى لا تتطلب مثل هذا النوع من التعامل كالأعمال الكتابية والهندسية • بل يمكن أن نقارن بين أداء الطلاب الذين يقومون بنشاط كبير خارج المنهج وأداء غيرهم من الطلاب الذين لا يقومون بهذه الأنشطة خلال فترة دراسية معينة • كما أننا كثيرا ما نلجأ الى الجماعات المهنية لحساب صدق اختبارات الميول (ومنها الميول المهنية) ومقاييس الاتجاهات• ومن الجماعات التى تستخدم فىحساب صدق مقاييس الاتجاهات الجماعات السياسية أو الدينية أو غيرها من الجماعات الخاصة التى تمثل وجهات نظر مختلفة اختلافا واضحا حول موضوعات اجتماعية معينة • وقد تستخدم جماعات معينة مثل الأحداث المنحرفين المودعين فى المؤسسات فى مقابل الأحداث غير المنحرفين ، أو العصابيين فى مقابل الأسوياء ، عند حساب صدق اختبارات الشخصية التىتتعلق بالتكيف الانفعالى أو الاجتماعى• وفى بعض هذه الاختبارات يلجأ الباحثون الى التشخيص الطبى النفسى باعتباره أساسا لاختيار الأسئلة ودليلا على صدق الاختيار ، الا أن التشخيص الطبى النفسى لابد أن تتوفر عدة شروط هامة ليصبح محكا مفيدا ، من أهمها أن يستند الى ملاحظة طويلة المدى ودراسة مفصلة لتاريخ الحالة ، أما اذا اعتمد على مجرد مقابلة شخصية قصيرة أو على فحص طبى نفسى غير عميق فلن يكون هناك سبب يدعونا الى توقع أن يكون مثل هذا التشخيص أفضل من الدرجة التى يحصل عليها الفرد فى الاختبار موضع البحث ، بعبارة أخرى

يصبح المحك المفترض (التشخيص الطبى النفسى) كالاختبار الجديد نفسه ، مجرد مؤشر يحتاج هو فى ذاته الى مزيد من البحث حول صدقه ولا يكون محكا كافيا نعتمد عليه فى تقدير صدق الاختبار ٠

٥ ـ التقديرات : سبق أن تحدثنا عن التقديرات ratings . التى يعطيها المعلمون لتلاميذهم أو المشرفون على العمال لعمالهم وغيرهم عند حديثنا عن المحكات الأخرى والتى نطلبها بهدف الحصول على معلومات تتصل بهذه المحكات مثل التحصيل المدرسى الأكاديمى أو الاداء فى نوع متخصص من الأعمال أو النجاح المهنى ٠ الا أن التقديرات قد تصبح محكا نعتمد عليه فى ذاته عند حساب الصدق حين تتطلب حكما شخصيا من جانب الملاحظ (المعلم أو المشرف) على الخاصية النفسية التى يقيسها الاختبار ٠ ومن ذلك أن نطلب من الملاحظين (المعلمين أو المشرفين) أن يعطوا تقديرات للمفحوصين فى خصائص معينة تقيسها الاختبارات مثل السيطرة أو القيادة أو الأصالة أو الابتكار أو الأمانة ٠٠ الخ ٠ هذه التقديرات تعد ذاتها محكات واستخدمها الباحثون فى حسابهم لصدق كثير من الاختبارات ، وبخاصة اختبارات الشخصية، وذلك لصعوبة الحصول على محكات موضوعية فى هذا الميدان وبخاصة بالنسبة للسمات الاجتماعية التى لا يمكن الحكم عليها الا من خلال مثل هذه التقديرات المعتمدة على العلاقة الشخصية ٠

والواقع أن التقديرات التى نحصل عليها من ملاحظين لديهم التدريب والخبرة وفى ظروف مضبوطة ضبطا تجريبيا دقيقا توفر لنا معلومات هامة عن المحك ٠ ومما يزيد ثقتنا فى هذا المحك أن نحصل على تقديرات مستقلة من أكثر من ملاحظ وحساب معاملات الموضوعية (راجع الفصل الرابع) حتى يمكن التغلب على العوامل الذاتية التى تؤثر فى ثبات هذه التقديرات ٠ كما يمكن أن تزداد دقة التقديرات اذا استخدمنا مقاييس التقدير rating scales بحيث تتضمن وحدات محددة تحديدا واضحا مع وجود ضمانات خاصة للتخلص من أخطاء التقدير الشائعة مثل أثر وقع الهالة halo effect (وهى عبارة عن ميل الملاحظين أو الفاحصين الى التأثر بسمة أو سمات غير مطلوبة ، وتكون فى صالح المفحوص أو فى غير صالحه وتؤدى الى التأثير فى أحكامهم وتقديراتهم للسمة موضع القياس) ، كما لابد أن تتوفر للمقدرين « ألفة بالسمات موضوع القياس » trait acquaintance وفهم لمعناها

بحيث يمكنهم الحكم على وجودها بدرجة معينة لدى المفحوصين · بعبارة أخرى لا يكفى أن يعرف الفاحص المفحوص فترة طويلة ليصبح قادرا على اصدار حكم عليه وانما يجب أن تكون قد تهيأت له فرصة ملاحظته فى مواقف مختلفة تظهر فيها السمة السلوكية موضوع الاهتمام · وفى جميع الحالات لا يجب أن يعطى الفاحصون تقديرات للمفحوص فى سمة ليس لديهم عنها معرفة كافية ·

وتلعب التقديرات دورا هاما فى مقاييس المحك الذى يعتمد على التقويم الكلينيكي · وفى مثل هذه الأحوال لابد من تدعيم الحكم الكلينيكى ببعض المعلومات الاضافية مثل التاريخ المفصل للحالة ودرجات الاختبارات الأخرى وغيرها من الوسائل الموضوعية، أما التقويم النهائى فيعتمد على عملية الحكم كما هو الحال فى جميع التقديرات ·

٦ ـ معاملات الارتباط بين الاختبار وغيره من الاختبارات :
من وسائل تحديد صدق الاختبار حساب معاملات الارتباط بين درجات الاختبار التجريبى (الجديد) وبين درجات اختبارات أخرى قائمة بالفعل تقيس ما يفترض الاختبار الجديد قياسه · وقد نلجأ الى هذه الطريقة اذا كان الاختبار الجديد صورة مختصرة أو مبسطة لاختبار موجود فعلا ، وفى هذه الحالة يكون الاختبار الأصلى محكا · كما يمكننا أن نقدر صدق اختبار لفظى (ورقة وقلم) باستخدام اختبار يدوى عملى مطول يستغرق وقتا طويلا وله معامل صدق مرتفع تم تحديده من قبل · وكذلك يمكن تقدير صدق اختبار جماعى باستخدام اختبار فردى (مثل اختبار ستانفورد ـ بينيه الذى استخدم كثيرا كمحك لحساب صدق الاختبارات الجماعية للذكاء) · وفى مثل هذه الأحوال يمكن اعتبار الاختبار الجديد فى أحسن الحالات صورة تقريبية للاختبار الأصلى · ويجب أن نلاحظ انه ما لم يكن الاختبار الجديد بديلا مبسطا أو مختصرا للاختبار الأصلى يصبح لا قيمة له ، لأنه لن يكون فى احسن حالاته الا تكرارا لاختبار موجود بالفعل ·

بعض مشكلات الصدق المرتبط بالمحكات : لعل من أهم المشكلات التى تعترض نظرية القياس النفسى بالنسبة للصدق المرتبط بالمحكات ما يتصل بمدى عمومية هذا النوع من الصدق ، وخاصة أن أهم ما يميزه أنه نوع من

الصدق العملى للاختبار في موقف معين بذاته ، كأن يحدد فالية الاختبـــار لبرنامج تدريبى معين ٠

وقد بدأت المشكلات في الظهور حين لوحظ أن معاملات الارتباط بين اختبارات الاستعدادات المقننة والأداء في اعمال يفترض فيها أن تتشابه مع ما تقيسه هذه الاختبارات ، وخاصة فى بحوث الصدق فى الأعمال الصناعية لوحظ أن معاملات الصدق اختلفت اختلافات بينة ، كما حصل الباحثون على ما يشبه ذلك في بحوث الصدق التى استخدمت التحصيل فى مختلف المقررات الدراسية محكا ، وأدى ذلك الى سيطرة اتجاه لدى مستخدمى الاختبارات حتى منتصف السبعينات خلاصته أن صدق الاختبار لا يقبل التعميم الى مواقف مختلفة ٠ وسادت فكرة « الخصوصية الموقفية » ، واعتبرت فى نفس الوقت عائقا لاستخدام الاختبارات المقننة فى انتقـــاء الأفراد ٠ الا أن الدراسات السيكومترية التى اجريت طوال العقدين الماضيين اكدت أن هذه الاختـلافات فى معاملات الصدق المرتبط بالمحكات هى «محض اصطناع احصائى» ناتج عن عن صغر حجم العينات المستخدمة فى حساب الصدق ، وعدم ثبات المحك ، وضيق مدىالفروق الفردية فى عينات التقنين ،وهذه جميعا تؤدى الى اختلاف معاملات الصدق المستخدمة فىحساب الصدق ، ودم ثبات المحك ، وضيق مدى الفروق الفردية فى عينات التقنين ، وهذه جميعا تؤدى الى اختلاف معاملات الارتباط من ناحية ، والى خفضها من ناحية أخرى ٠

وقد امكن التغلب على هذه المشكلة فى السنوات الأخيرة من خلال نموذج جديد اقترحه شمدت Schmidt وهانتر Hunter واطلقا عليه اسم «النموذج البايزىلتعميم الصدق» Bayesian Validity Generalization Model وطبقاه على عينات كثيرة اختيرت من تخصصات مهنية كثيرة واكد أن صدق اختبارات الاستعدادات اللفظية والعددية والاستدلالية يمكن أن يعمم الى عدد كبير من المهن ، وان الاختلافات السابقة لم تكن الا محض « مصادفة » ٠

(هـ) صدق التكوين الفرضى :

يقصد بصدق التكوين الفرضى Construct validity مدى قيـــاس الاختبار لتكوين فرضى أو مفهوم نفسى أو سمة ٠ ومن أمثلة هذه التكوينات

الفرضية الذكاء والفهم الميكانيكى والطـــلاقة اللغوية والعصـــابية والقلق • وغيرها • وكل من هذه المفاهيم أو التكوينات يستنتج من أساليب الأداء كما يفسر حدوث الاتساق الملحوظ فى الاستجابة (٢١) • ويعتمد هذا النوع من الصدق على وصف أوسع ويتطلب معلومات اكثر حول الخاصية السلوكية موضوع القياس نحصل عليها من مصادر مختلفة • وأى معلومات تفيد فى القاء الضوء على المفهوم الذى نحاول قياسه وعلى العوامل التى تؤثر فيـه تضاف الى رصيد صدق الكوين الفرضى •

وتوجد أساليب كثيرة يمكن أن تسهم فى تحديد هذا النوع من الصدق، وهى ما يأتى :

١ ـ **تمايز العمر والتغيرات الارتقائية** : يمكن أن نلجأ الى هذه الطريقة فى حسابنا لصدق اختبارات ذكاء الأطفال (كاختبار ستانفرد ـ بينيه ، أو معظم اختبارات الذكاء لأطفال ما قبل المدرسة) ، وذلك لتحديد ما اذا كانت الدرجات تتزايد بتقدم العمر مفترضين أن القدرات تنمو مع العمر خـــلال الطفولة وبالتالى نتوقع زيادة فى درجات الاختبار اذا كان الاختبار صادقا • ويعود الفضل الى بينيه فى صياغة هذا المفهوم الخاص بتحديد المقياس العمرى للذكاء •

ومن الواضح أن محك تمايز العمر لا يمكن استخدامه فى الوظائف السلوكية التى لا تظهر تغيرات مطردة مع تقدم العمـــــر ومن ذلك مقاييس الشخصية بوجه عام • كما أننا يجب أن نشير الى أن هذا المحك محك سلبى وليس محكا ايجابيا • بعبارة أخرى اذا لم تستطع درجات الاختبار أن تتزايد مع تزايد العمر فان ذلك قد يدل على أن الاختبار غير صادق فى قياسه للقدرات التى يدعى أنه عينة لها ، الا أن تحسين الاختبار بحيث يصبح مقياسا متزايدا مع تزايد العمر لا يحدد لنا فى ذاته الميدان الذى يقيسه الاختبار • فمثلا يظهر كل من الطول والوزن زيادة منتظمة مع تزايد العمر ومع ذلك لا يمكن أن نعتبر أحدهما أو كليهما من مقاييس الذكاء • وكذلك يجب أن نؤكد أنه فى الاختبار النفسى الذى يقيس خصائص سلوكية تتزايد مع العمر انما تتأثر هذه الزيادة أو النمو بظروف البيئة التى يقنن فيها الاختبار ، فالظروف البيئية والحضارية قد تستثير بعض الخصائص السلوكية وقد تعطل البعض الآخر • ومعنى ذلك

أن محك تمايز العمر لا يمكن أن نعتبره محكا شموليا بحال من الأحوال ، وانما هو كغيره من المحكات محكوم بالموقف الحضارى الخاص الذى نشتقه منه ٠

٢ ـ **معاملات الارتباط بالاختبارات الاخرى :** وقد نلجأ فى تقدير صدق التكوين الفرضى الى حساب معاملات الارتباط بين الاختبار التجريبى وبين الاختبارات التى تتشابه معه نظريا فى العمليات التى يقيسها ٠ ويسمى كامبل Campbell وفيسك Fiske الصدق التقاربى converant validity ومن امثلة ذلك معامل ارتباط اختبار الاستدلال الكمى بدرجات المفحوصين فى اختبار التحصيل الرياضى ٠ الا أن هذا النوع من معاملات الارتباط لا يكفى ولذلك يقترح كامبل وفسك ضرورة أن يضاف الى ذلك حساب معاملات الارتباط بين الاختبار واختبارات أخرى يفترض نظريا أنها تختلف عنه ، ويسميان هذا النوع من الصدق باسم الصدق التمييزى discriminant validity

ويستخدم الصدق التمييزى لمعرفة مدى تحرر الاختبار التجريبى من أثر العوامل التى لا علاقة لها بالعملية السلوكية موضع القياس ، فمثلا اذا حصلنا على معامل ارتباط صفرى بين اختبار معين للاستعدادات واختبارات الفهم اللغوى فان ذلك قد يعنى أن هذا الفهم اللغوى لا يرتبط بالاستعداد موضوع القياس فى الاختبار الجديد وكثيرا ما تتضمن كتيبات الاختبارات معاملات الارتباط المنخفضة بين الاختبار التجريبى واختبارات أخرى تقيس عمليات مختلفة وهى مؤشرات على الرغم من أنها غير مباشرة وسلبية الا انها تدل فى نفس الوقت على صدق الاختبار ، لأنه اذا كانت هذه المعاملات مرتفعة كان ذلك مدعاة للشك فى صدق الاختبار اذ تعنى أن الاختبار يقيس عمليات أخرى لم توضع فى الحسبان ٠ الا ان معاملات الارتباط المنخفضة هذه قد لا تعنى فى ذاتها أن الاختبار التجريبى صادق ٠ ولكى تستخدم طريقة الصدق التقاربى والصدق التمييزى بكفاءة يقترح كامبل وفسك تصميما تجريبيا خاصا يسمياه « مصفوفة السمات المتعددة والطرق المتعددة » Multitrait-Multimethod Matrix

ويتطلب هذا التصميم تقويم سمتين نفسيتين أو اكثر بطريقتين مختلفتين أو اكثر ، ثم تحسب جميع معاملات الارتباط الممكنة بين درجات الاختبارات

التى تقيس السمات موضع البحث · ولكى نوضح هذه الطريقة نعطى المثال الافتراضى التالى الذى يقترحه الباحثان ·

نفوض أننا أردنا قياس ثلاث سمات نفسية هى (أ) السيطرة ، (ب) الاجتماعية ، (حـ) دافعية الانجاز ، بثلاث طرق مختلفة هى (١) قائمة من نوع التقرير الذاتى ، (٢) أسلوب اسقاطى ، (٣) تقديرات الأقران والزملاء ، فاننا نشير فى هذه الحالة الى سمة السيطرة كما تقاس بالتقرير الذاتى بالرمز ($أ_1$) وكما تقاس بالاختبار الاسقاطى ($أ_2$) ، والى سمة دافعية الانجاز كما تقاس بتقديرات الأقران بالرمز ($جـ_3$) ، وهكذا بالنسبة لجميع البدائل الممكنة ·

ويوضح الجدول رقم (٥) مصفوفة ارتباط جميع الدرجات التى يمكن الحصول عليها بهذه الطريقة ، وقد وضعت معاملات الثبات الخاصة بطريقة قياس كل سمة بين قوسين ·

وفى هذا الجدول نجد (١) معاملات الصدق فى صورة علاقات بين الدرجات التى حصل عليها المفحوصون لنفس السمة بطرق مختلفة مستقلة (كما هو مفترض فى اجراءات الصدق المعتادة) ، (٢) معاملات الارتباط بين مختلف السمات التى قيست بنفس الطريقة ، و (٣) معاملات الارتباط بين السمات المختلفة التى قيست بطرق مختلفة · ولكى يحصل الباحث على صدق التكوين الفرضى لابد بالطبع أن تكون معاملات الصدق (النوع الأول) أعلى من النوعين الآخرين من معاملات الارتباط · فنحن نتوقع مثلا أن معامل الارتباط بين دافعية الانجاز كما تقاس بوسائل التقرير الذاتى وكما تقاس بالاختبار الاسقاطى يكون أعلى من معامل الارتباط بين دافعية الانجاز والسيطرة كما تقاس بوسائل التقرير الذاتى ، فاذا حدث أن العامل الأخير كان مرتفعا فاننا قد نستنتج من ذلك أن أداء الفرد فى الاستخبار أو القائمة المستخدمة يتأثر بعامل مشترك لا يرتبط بما يقيسه (أى نوع من تباين الخطأ الذى تناولناه بالتفصيل فى الفصل الرابع المخصص لموضوع الثبات) ، وقد يكون هذا العامل متصلا بالفهم اللغوى للاسئلة أو العبارات ، أو رغبة المفحوص فى أن يظهر بصورة موجبة فى جميع السمات (المرغوبية الاجتماعية) ·

وتذكر أنستازى (١١ : ١٤٩) أن معامل الثبات فى مصفوفة السمات

جدول رقم (٥) مصفوفة ارتباط متعددة وطرق متعددة

(عن ١١ : ١٤٨)

السمات		الطريقة ١ أ١	الطريقة ١ ب١	الطريقة ١ حـ١	الطريقة ٢ أ٢	الطريقة ٢ ب٢	الطريقة ٢ حـ٢	الطريقة ٣ أ٣	الطريقة ٣ ب٣	الطريقة ٣ حـ٣
الطريقة ١	أ١	(٨٩ر)								
	ب١	٥١ر	(٨٩ر)							
	حـ١	٣٨ر	٣٧ر	(٧٦ر)						
الطريقة ٢	أ٢	٥٧ر	٢٢ر	٠٩ر	(٩٣ر)					
	ب٢	٢٢ر	٥٧ر	١٠ر	٦٨ر	(٩٤ر)				
	حـ٢	١١ر	١١ر	٤٦ر	٥٩ر	٥٨ر	(٨٤ر)			
الطريقة ٣	أ٣	٥٦ر	٢٢ر	١١ر	٦٧ر	٤٢ر	٣٣ر	(٩٤ر)		
	ب٣	٢٣ر	٥٨ر	١٢ر	٤٣ر	٦٦ر	٣٤ر	٦٧ر	(٩٢ر)	
	حـ٣	١١ر	١١ر	٤٥ر	٣٤ر	٣٢ر	٥٨ر	٥٨ر	٦٠ر	(٨٥ر)

المتعددة والطرق المتعددة يمثل درجة الاتفاق بين مقياسين لنفس السمة بينهما اقصى حد من التشابه (كما هو الحال فى الصور المتكافئة لنفس الاختبار) ، اما الصدق فيمثل درجة الاتفاق بين مقياسين لنفس السمة بينهما أقصى حد من الاختلاف (كما هو الحال فى درجات الاختبار وتقديرات المعلمين أو الرؤساء)٠ وطالما أن التشابه والاختلاف بين الطرق هو مسألة نسبية فاننا - من الوجهة النظرية البحتة - يمكن أن نعتبر الثبات والصدق يقعان على مدى متصل واحد ٠ ولعلنا نؤكد فكرة هذا المتصل بين الثبات والصدق اذا ذكرنا أن الثبات فى موجزه يهتم بتباين الخطأ ، بينما الصدق يكاد يتطابق مع مفهوم التباين الحقيقى ، وكلاهما يؤلف التباين الكلى أو الفروق الفردية فى الدرجات التى يحصل عليها المفحوصون فى الاختبار ٠

ويقترح فسك منذ مطلع السبعينات أن يقوم الباحثون فى مجال الاختبارات النفسية بتحليل مصفوفات عديدة لمعاملات الارتباط بين مقاييس عدد من التكوينات الفرضية (السمات) لطرق مختلفة ، ومن فروضه الأساسية أننا نتوقع فى جميع الحالات أن تكون مقاييس نفس السمة تعطينا نمطا لمعاملات الارتباط يتشابه مع مقاييس سمة أخرى ٠ فمثلا معامل الارتباط بين تقديرات الأقران للسيطرة واستخبار التقرير الذاتى لهذه السمة يظهر على وجه التقريب نفس نمط معاملات الارتباط بين نفس الطريقتين فى قياس سمات المثابرة أو الانتساب أو الاستقلال أو أى سمة أخرى تستخدم فى قياسها هاتان الطريقتان ٠ وقد اختبر هذا الفرض الذى يقترحه فسك فى عدد من البحوث التى أجريت طوال العقدين الماضيين فى مجال مقاييس الشخصية وأظهرت النتائج أن معاملات الارتباط بين السمات فى هذه الحالة لا تظهر الا تشابها جزئيا مع وجود قدر من الفروق الدالة يرجع الى اختلاف الطرق (التباين بين الطرق) بل قد نحصل على أنماط مختلفة لمعاملات الارتباط حتى داخل الطرق (التباين داخل الطرق) ٠ وقد وجدت هذه الفروق مثلا بين وسائل التقرير الذاتى التى بنيت مستقلة واختلفت فى القالب (صواب - خطأ - أو مقارنة ثنائية او قائمة عبارات) أو فى محتوى المفردات المستخدمة فى الاختبار ٠ ومع ذلكفقد يوجد قدر دال من التباين يمكن قياسه بالطرق المختلفة ويشير الى تكوين فرضى له معنى سيكولوجى ويمكن تعميمه عبر الطرق المختلفة المستخدمة فى تقويمه نفسيا ، وفى هذا وحده معنى صدق التكوين الفرضى ٠

٣ - **التحليل العاملي** : قد يكون التحليل العاملي من أهم الوسائل التي تستخدمها في تحديد صدق التكوين الفرضي ، وهو في جوهره أسلوب لتحديد بنية السمات النفسية ، وتقوم على تحليل العلاقات بين البيانات الامبريقية كما تتمثل هذه العلاقات في صورة معاملات الارتباط. فنقطة البدء في التحليل العاملي حساب معاملات الارتباط بين درجات كل اختبار والاختبارات الأخرى • التي يجب أن تكون قد طبقت على نفس العينة من المفحوصين • وعندما نفحص الجدول الذي يتضمن هذه المعاملات (والذي يسمى فنيا مصفوفة الارتباط) ونطبق عليه طرقا احصائية خاصة لا يتسع المقام لشرحها قد نلاحظ وجود تجمعات معينة بين الاختبارات مما يوحي بوجود سمات مشتركة بينها • فمثلا اذا كانت الاختبارات التي يرتبط بعضها ببعض ارتباطا عاليا من نوع اختبارات المفردات والمشابهات والأضداد وتكملة الجمل وكانت معاملات ارتباط هذه الاختبارات بالاختبارات الأخرى منخفضة نستنتج من ذلك وجود عامل احصائي تتشبع به المجموعة الأولى من الاختبارات تشبعا عاليا وتتشبع به المجموعة الثانية تشبعا منخفضا (والتشبع مفهوم احصائي يستخدم في التحليل العاملي ليدل على علاقة المتغير بالعامل)•

وبعد ذلك نبدأ عملية تفسير العامل الاحصائي الذي تم الحصول عليه فاذا كانت الاختبارات من نوع الأداء الاقصى كما هو في مثالنا الحالي يفسر العامل بأنه « القدرة على الفهم اللغوي » وقد يفسر العامل بأنه « مهارة » اذا كانت الاختبارات من النوع الذي يسمى « نفسحركي » ، وقد يفسر بأنه « سمة » أو « ميل » أو « اتجاه » أو « قيمة » الخ حسب طبيعة الاختبارات التي تنتمي الى الأداء المميز •

وبالطبع نحن لا نعتمد في هذا التحليل على مجرد تأمل مصفوفة الارتباط والا كانت العملية صعبة ومعقدة وغير مؤكدة بل مستحيلة وخاصة اذا كانت بطارية الاختبارات الكبيرة • ولذلك تستخدم الأساليب الاحصائية الأكثر دقة لتحديد العوامل المشتركة التي يمكن اعتبارها مسئولة عن معاملات الارتباط الناتجة •

وفي التحليل العاملي نلاحظ أن عدد المتغيرات التي نصف في ضوئها أداء كل فرد تختصر باستمرار من عدد كبير من الاختبارات الأصلية الى

عدد أصغر نسبيا من العوامل أو السمات المشتركة · ثم يمكن وصف الفرد فى ضوء درجاته فى العوامل القليلة العدد · بعبارة أخرى يمكن القول أن من أهم أغراض التحليل العاملى تبسيط وصف السلوك باختزال عدد المتغيرات اى انه يسير فى اتجاه العلم الصحيح من التعدد والكثرة التى تتمثل فى متغيرات الاختبارات الى القلة التى تتمثل فى العوامل المشتركة او السمات · الا أن للتحليل العاملى وظيفة أخرى هى فى رأينا أكثر ارتباطا بمفهوم صدق التكوين الفرضى وهى محاولة اختبار صحة فروض معينة حول البنية العاملية للاختبار الواحد من خلال تحليل مصفوفة ارتباط مفرداته أو أسئلة أو بطارية من الاختبارات النفسية ، وهى وظيفة اشار اليها فؤاد ابو حطب (٩) منذ مطلع السبعينيات ثم ازداد الاهتمام بها فى الوقت الحاضر فيما يسمى التحليل العاملى المدعم Confirmatory Factor Analysis .

وبعد تحديد العوامل يمكن أن نفيد منها فى وصف التكوين العاملى للاختبار · فكل اختبار يمكن أن يوصف فى ضوء العوامل الأساسية التى تحدد درجاته بالاضافة الى تشبعه بكل عامل · وتشبعات العوامل هذه تمثل بدورها معاملات ارتباط بين الاختبار والعامل · وهذا العامل يسمى الصدق العاملى الذى هو فى الواقع ماعمل ارتباط بين الاختبار وبين ما هو شائع أو مشترك أو عام فى مجموعة الاختبارات أو غيرها من المؤشرات الدالة على السلوك · ومن المؤكد أن المتغيرات التى نقوم بتحليلها عامليا تتضمن معلومات نشتقها من الاختبارات أو غيرها من أدوات التقويم النفسى كالتقديرات والاحكام ، وهذه جميعا يمكن أن تستخدم فى محاولة استكشاف الصدق العاملى لاختبار معين وتحديد السمات المشتركة التى يقيسها ، كما يمكن أن يطبق التحليل العاملى على مصفوفات ارتباط المفردات التى يتالف منها الاختبار الواحد ·

٥ ـ **الاجراءات التجريبية** : من الوسائل الاخرى التى يمكن ان نلجأ اليها لتحديد صدق التكوين الفرضى ما توفره لنا الاجراءات التجريبية من معلومات حول اثر متغيرات معينة على درجات الاختبار · فاذا كان الاختبار التجريبى (الجديد) معدا لقياس القابلية للقلق مثلا يمكن ان نطبق هذا الاختبار على مجموعة من الأفراد توضع فى موقف يثير القلق (مثل الامتحان فى ظروف ضاغطة) ثم يحسب معامل الارتباط بين درجات الأفراد فى اختبار

جدولا رقم (٥) يوضح أنواع الصدق واستعمالات كل منها

نوع الصدق	السؤال المطروح	الاجراء	الاستخدام الرئيسي	أمثلة
صدق المحتوى	هل اكتسب المفحوص مقدار كافيا من المعلومات أو المهارة من البرنامج التدريبي الذي يتعرض له ؟	نقارن بين أسئلة الاختبار من الوجهة المنطقية بالمحتوى الذي نفترض أن أسئلة الاختبار تنتمي اليه .	اختبارات التحصيل المدرسي والكفاءة المهنية .	فحص اختبار القدرة على الاختزال لمعرفة هل يحتوى مضمونه بالفعل على مهارات العمل المكتبي .
الصدق المرتبط بالمحكات (١) الصدق التنبؤى	هل تتنبأ درجات الاختبار بأداء معين له أهمية في المستقبل ؟ .	اعطاء الاختبار ثم نحصل على مقياس لمحك بعد فترة زمنية ثم نقارن الاختبار بالمحك	الاختبارات التي تستخدم في الانتقاء والتصنيف .	مقارنة اختبارات القبول بالكليات أو المعاهد بتحصيل الطلاب فيما بعد .
الصدق المرتبط بالمحكات (٢) الصدق التلازمي .	هل تشخص درجات الاختبار صعوبات الأداء في الوقت الحاضر ؟ .	نعطى الاختبار ثم نحصل على مقياس مباشر دون فاصل زمني ثم نقارن الاختبارات بالمحك .	الاختبارات التي تستخدم في أغراض التشخيص التربوى أو الكلينيكي .	مقارنة أداء المتخلفين عقليا بالأسوياء في اختبار للذكاء .
صدق التكوين الفرضي	ما المعنى النفسي للمفهوم الذي يقيسه الاختبار ؟	صياغة الفروض حول الاختبار ثم التحقق منها بالاجراء التجريبي أو الاحصائي المناسب .	الاختبارات التي تستخدم في أغراض البحث العلمي حول الاختبارات النفسية .	دراسة اختبار القدرة الموسيقية لتحديد الى أي حد تتأثر الدرجات فيه بعوامل التدريب الموسيقي .

القلق قبل التجريب وبين مختلف المؤشرات الفسيولوجية وغيرها التى تدل على القلق خلال التجريب (موقف الامتحان) وبعده • ويمكن القول ان هذا الاجراء التجريبي يشبه الى حد كبير ما نلجأ اليه فى تقدير الصدق التنبؤى• باستثناء الدور الايجابى للباحث فى تقديم مثير القلق والتحكم فيه • اما المقاييس المباشرة الدالة على القلق فيمكن اعتبارها محكا نراجع عليه تنبؤاتنا من درجات الاختبار •

وقد تفيد الوسائل التجريبية أيضا فى تحقيق الفروض المختلفة حول الاختبارات ، ففى مثالنا السابق ذكره يمكن تطبيق اختبار القلق قبل التجريب (اعطاء خبرة القلق) وبعده ثم نتحقق مما اذا كانت درجات الاختبار ترتفع ارتفاعا له دلالة عند اعادة الاختبار وبالطبع يمكن اســـتخدام تصميمات تجريبية تعتمد على المجموعات الضابطة والتجريبية ، فاذا كانت النتيجــة ايجابية فان ذلك يدل على أن درجات الاختبار تعكس المستوى الراهن للقلق، وتبرهن على صدق التكوين الفرضى الذى يقيسه •

العوامل المؤثرة فى الصدق :

اذا صيغ الصدق فى صورة معامل ارتباط فانه يصبح مثل معامل الثبات يتأثر بالعوامل المختلفة التى تؤثر فى معامل الارتباط وخاصة ما يتصــل بتجانس العينات والتى تناولناها بالتفصيل فى الفصل السابق الا أن هناك عوامل تؤثر فى الصدق تأثيرا خاصا ولابد من أن توضع فى الاعتبار •

١ ـ **طبيعة عينة التقنين** : من المهم لفهم طبيعة معامل الصدق وصف طبيعة عينة التقنين وصفا مفصلا ، لأن الاختبار الواحد قد يقيس عمليات نفسية مختلفة اذا طبق على عينات تختلف فى العمر او الجنس أو المهنة أو المستوى التعليمى أو أى متغير آخر يرتبط بمفهوم الصدق • وعلى ذلك فان الاختبار الواحد قد يكون له صدق مرتفع فى التنبؤ بمحك معين فى بعض العينات ، بينما ينخفض صدقه بالنسبة الى عينات اخرى وربما يكون مقياسا صادقا لعمليات نفسية مختلفة فى عينتين مختلفتين ، ومن ذلك مثلا اختبار العمليات الحسابية الذى قد يكون مقياسا صادقا للاستدلال الحسابى عند أطفال الصفوف الدنيا فى المرحلة الابتدائية ، بينما يقيس السرعة العددية بصدق عند طلاب المرحلة الثانوية •

٢ - **طبيعة العلاقة بين الاختبار والمحك** : من المعروف أن حساب معامل الارتباط بين الاختبار والمحك ، اذا استخدمت فيه طريقة كارل بيرسون أو تقريباتها المختلفة ، فان ذلك يتضمن افتراضا جوهريا وهو أن العلاقة بين المتغيرين خطية linear على امتداد مدى الفروق الفـــردية فيهما ٠ وبالطبع فان الافتراض قد يكون صحيحا فى عدد كبير من الحالات ، الا انه فى بعض الحالات قد لا تكون العلاقة كذلك ، ومن ذلك مثلا حين يتطلب أحـد الأعمال حدا أدنى من فهم القراء بعين العاملين فيه على قراءة بعض التعليمات وفى هذه الحالة نجد أنه حالما توافر هذا الحد الأدنى فان أى زيادة عليه فى القدرة على فهم القراءة يصبح غير مرتبط بدرجة النجاح فى العمل ٠ وتوصف العلاقة بين الاختبار والمحك فى هذه الحالة بأنها غير خطية non-linear ومن أمثلة ذلك العلاقة المنحنية curvilinear والمثلثة triangle أو العـــــلاقة التى تدل على اختلاف التباين heteroscedasticity ، ومن أمثلة النوع الأخير من العلاقات حين نجد أن الأداء فى اختبار الذكاء يكون شرطا ضروريا ولكنه ليس كافيا للنجاح المدرسى ٠ وفى هذه الحالة نجد التلاميذ من ذوى الذكاء المنخفض يؤدون أداء سيئا فى المدرسة ، ولكننا حين نفحص أداء التلاميذ من ذوى الذكاء المرتفع سوف نجد أن بعضهم يؤدون أداء جيدا فى المدرسة وبعضهم الآخر قد يؤدى أداء سيئا لأسباب تتصل بنقص الدافعية ، أو فقدان الميل أو غيرها ، وفى هذه الحالة نجد تباينا فى التحصيل بين ذوى الدرجات المرتفعة فى الذكاء أكثر اتساعا مما نجــــده بين ذوى الدرجات المنخفضة فيه ٠

٣ - **اثر ثبات الاختبار فى صدقه** : من بديهيات القيـــاس النفسى أن الاختبار غير الثابت وغير الدقيق لا يمكن أن يكون صــــادقا ٠ لأن معنى انخفاض ثبات الاختبار أن تباين الخطأ فيه أكبر من التباين الحقيقى ، وبالطبع فان تباين الخطأ فى درجات المفحوصين لا يمكن أن يرتبط ارتباطا دالا (أى يتجاوز مستوى الصدفه) مع أى محك ٠ وعلى هذا كلما زاد تباين الخطأ يقل معامل الصدق ، لأن الصدق هو التباين الحقيقى للاختبار أما تحرر الاختبار تماما من تباين الخطأ فان يصل بالثبات الى حد التمام ويصبح التباين الكلى فى هذه الحالة مساويا للتباين الحقيقى، وفى هذه الحالة النظرية المستحيلة الحدوث وحدها يمكن للثبات أن يساوى الصدق ٠

الا أن الاختبار النفسى - كغيره من صور القياس - يكون فى العادة مشوبا بقدر من تباين الخطأ ، ولهذا فان من المتوقع فى هذه الحالة الا يزيد معامل صدقه عن معامل ثباته ، لأنه لا يكون منطقيا على أى نحو أن يرتبط الاختبار مع نفسه أو مع ما يماثله أو يشبهه من صور متكافئة (وهو معنى الثبات) ارتباطا أقل من ارتباطه مع اختبارات أو مقاييس أخرى مختلفة عنه هى ما يسمى المحكات (وهو معنى الصدق) ، ولذلك فكثيرا ما يقـــال أن الحدود القصوى لمعاملات الصدق المحسوبة هى معاملات الثبات والتى عادة ما لا تتجاوزها النتائج الامبريقية للاختبارات .

الا أن هذه العلاقة بين الثبات والصدق قد أسىء عرضها فى عدد من البحوث النفسية التى أجريت فى مصر والعالم العربى فى السنوات الأخيرة ، وخلاصة هذا الخطأ أن بعض الكتب الاحصائية باللغة العربية أشارت الى مفهـــوم مؤشر الثبـــات index of reliability وهو عبارة عن الجذر التربيعى لمعامل الثبات بمصطلح الصدق الذاتى ، وهى تسمية غير دقيقة فى رأى كاتب هذا الفصل ، لأنها تحمل المؤشر مالايعنيه ، فهو لا ينتمى الى الصدق بمعناه العلمى الدقيق .

وتلقف باحثونا المستعجلون الفكرة ووجدوا فيها تعويضا عن عجزهم عن البحث فى الطرق الصحيحة لتحديد صدق الاختبارات ، فنجد الواحد منهم يحسب ثبات اختباره ثم يحسب الجذر التربيعى لهـذا الثبات ويشـــير اليه بالصدق ، وكفى الله المؤمنين شر القتال . لون آخر من الوان العبث العلمى الذى يشهده المسرح المعاصر لعلم النفس .

وخلاصة فكرة مفهوم مؤشر الثبات (أو ما يشار اليه خطأ بالصـــدق الذاتى) أن مربع معامل للصدق (أى مربع معامل الارتباط بين الاختبار والمحك) يجب أن يكون أقل من معامل الثبات ، وذلك لأن معامل الثبات يساوى مربع معامل الارتباط بين الاختبار ومتوسط الأصل الاحصائى العام الذى تشتق منه العينة . ومعنى ذلك أن مربع معامل الصدق يجب أن يكون أقل من مربع معامل ارتباط الاختبار بمتوسط الأصل الاحصائى ، وبعبارة أخرى أن الاختبار لا يمكن أن يرتبط بمتغير خارجى (وهو المحك) أعلى من ارتباطه بمتوسط أصله الاحصائى العام .

لتوضيح هذه الفكرة نعطى المثال التالى : نفرض أن معامل ثبات أحد اختبارات الذكاء بلغ ٨٧ر ، معنى ذلك أننا لو استبعدنا أخطاء العينة ، فان هذا الاختبار لا يمكن ان يتنبأ بالتحصيل المدرسى (وهو المحك) بمربع معامل صدق يزيد عن هذا المقدار (أى ٨٧ر) ، وعلى ذلك فان معامل الصدق لا يجب ان يزيد بحال من الأحوال عن الجذر التربيعى لمعامل الثبات (وهو فى هذه الحالة ٩٣ر٠) ٠ ان هذا الجذر التربيعى ليس معامل صدق) وانما يدل على الحد الذى لا يمكن أن تتجاوزه معاملات الصدق ٠ أنه قد يعنى أن معامل الصدق قد يكون أعلى من معامل الثبات ـ فى ضوء هذا الحد الأقصى ـ الا ان هذا الافتراض النظرى لم يتحقق تجـــريبيا ابدا ، وجميع البيانات الامبريقية التى يجمعها الباحثون بنزاهة تتفق مع البديهية التى دارت حولها مناقشتنا السابقة وهى أن معاملات الصدق لابد ان تكون أقل من معاملات الثبـــات ٠

الفصـــل السادس

تصحيح الاختبارات وتفسير الدرجات

تصحيح الاختبارات :

يعد تصحيح الاختبارات باعطاء الفرد درجة او تقديرا ، وتفسير هذه الدرجة هو خطوة هامة ، ولو أنه فى ذاته يعد مقدمة لاتخاذ قرار عملى أو تفسير علمى عن الفرد أو مجموعة الأفراد موضوع القياس • وعلى كل فان مشكلات تصحيح الاختبارات لا تقل أهمية وخطورة عن المشكلات الأخرى التى أثرناها فى هذا الباب ولعل أخطر هذه المشكلة ما يتصل بموضوعية التصحيح• والموضوعية objectivity هى مشكلة العلم عامة ، ومشكلة العلوم الانسانية خاصة ، وقد تناولها فى السنوات الأخيرة عدد من الفلاسفة وعلماء النفس (٢ ، ٣ ، ٤ ، ٥) ، ودون ان ندخل فى تفاصيل مناقشـــة شائكة حول هذه المسألة ، نشير هنا الى اننا نتناول الموضوعية فى اطار محدد وهو تصحيح الاختبارات• صحيح أن المسالة تتداخل مع جميع مراحل بناء الاختبار وتطبيقه وتقنينه ، وكلها تتصل بدور الباحث أو الفاحص أو المصحح •

وبالطبع فان المعنى المباشر للموضوعية هو تحديدها « سلبيا » ولهذا كثيرا ما نجدها تعرف بتحرر الفاحص من العوامل الذاتية كالتحيز والتعصب والآراء المسبقة والافكار الجاهزة التى يحاول فرضها على الوقائع ، الا أن هذا التحديد « السلبى » ليس تعريفا مفيدا للمفهوم ، وكل ما يعطيه لنا من معنى هو أن الموضوعية مقابل الذاتية • وفى رأى كاتب هذا الفصل أن هذه المقابلة الثنائية ليست هى التناول الصحيح للمسـالة ، فلغة الأضداد ليست هى اللغة الصحيحة للعلم ، ولا لمنهجه والأصح أن نتصورها على هيئة متصل يمتد على خط مستقيم يبدأ بالموضوعية الكاملة (شبه الآلية الميكانيكية) وينتهى بالذاتية الكاملة (الاعتماد الكامل على الهوى والغرض) وبينهما توجد درجات كالمسمية عند سيد عثمان (٣) « البصيرة العلمية » التى تتأتى بالخبرة •

وعلى وجه العموم نحن في‍ه حاجة فى مجال الاختبارات النفسية أن تكون عمليات تطبيق الاختبار وتصحيحه وتفسير درجاته مستقلة نسبيا عن الحكم الشخصى للفاحص ، حتى ولو توافر له قدر كاف من هذه البصـــيرة العلمية ، ولكى نتأكد من ذلك فاننا عادة ما نلجأ الى تحديد درجة « اتفاق الملاحظات والأحكام اتفاقا مستقلا » وهذا هو التعريف الذى يقترحه كاتب هذا الفصل للموضوعية · وهو تعريف موجب يتجاوز التعريفات «السالبة» المعتادة، بالاضافة الى أنه قابل للتناول والمعالجة · فهذا « الاتفاق المستقل » فى الملاحظات والأحكام والتقديرات يمكن تحديده بحساب معامل الارتباط بين عدد من الفاحصين فى ملاحظتهم أو حكمهم أو تقويمهم ـ مستقلين بعضهم عن بعض ـ لنفس المفحوصين ، وفى هذه الحالة تصبح « الأهواء الذاتية » للفاحصين جزءا من تباين الخطأ ، ويدخل الأمر كله فى قضية ثبات الاختبار وهو نوع من الثبات وثيق الصلة بموضوعنا الراهن ·

كما يمكن تحديد موضوعية الفاحص الواحد « بطريقة مماثلة وذلك بحساب معامل الارتباط بين عدد من ملاحظاته أو أحكامه او تقديراته أو تقويماته لنفس المفحوصين والتى تختار كعينات من سلوكه ، ومن المؤكد أن هذا المعامل اذا كان منخفضا فانه يدل على « ذاتية المصحح » أما اذا كان مرتفعا فقد يكون مؤشرا على « البصيرة العلمية » لديه ، بشرط أن نتأكد من عدم تداخل ملاحظاته بعضها مع بعض ، والا كان معامل الارتباط المرتفع هذا مؤشرا زائفا على هذه البصيرة ، فقد يدل فى هذه الحالة على اتســـــاق التحيزات أو الاهواء الخاصـــة وفى جميع الحالات يجب التنبـــه الى أن الموضوعية مسألة درجة ، وان كانت أكثر وضوحا فى نمط الاختبارات الذى نسميه « امتحانات المقال » · ففى هذا النوع كثيرا ما يتساءل الشخص لماذا حصل على درجة منخفضة · ومن المؤكد أن السبب الرئيسى هو أنه فى هذا النوع من الامتحانات يكون من الصعب تحديد الاجابة الجيدة وبالتالى يكون من الصعب اعطاء درجة مناسبة لاستجابة ليست صحيحة صحة كاملة ·

وبالطبع لا يمكن للباحث النفسى أن يجرى دراسة للسلوك كما لا يمكن ان يتخذ قرارا عمليا اذا لم تكن معايير التصحيح على درجة من الموضوعية، ولذلك يقترح البعض لحل هذه المشكلة اما تحديد قواعد للحكم يلتزم بهـــا المصححون للاختبــــارات التى تتطلب اجابات حرة او اســـتخدام الأسئلة

الموضوعية التى تتطلب التعرف على الاجابة الصحيحة من بين عدة اجابات، ويستخدم ، تصحيحها مفاتيح معدة مقدما ٠

تصحيح الأسئلة التي تتطلب استجابات حرة :

لا يزال القياس الفردى يستخدم أسئلة ومشكلات تتطلب درجة من الحكم اثناء التصحيح ، ومع ذلك فمن الممكن اعداد بعض الوسائل التى تتوفر فيها درجة من الموضوعية فى هذا التصحيح وبخاصة فيما يتعلق بالخصائص الرئيسية للسلوك ٠ فعند تقويم سلوك معين مثل الخط ، أمكن للباحثين اعداد نموذج للتصحيح يحتوى على عينات من الخطوط تمثل مستويات مختلفة من الجودة وعلى المدرس أن يحدد العينة الأكثر شبهابخط التلميذ حتى يحدد الدرجة التى يعطيها له ٠ وكذلك يمكن اعداد مقاييس معتمدة على تقدير الانتاج للحكم على « نوع » بعض ألوان السلوك الصناعى أو التجارى أو الادارى ٠ ومع أن تحديد قواعد تصحيح بعض الاختبارات التى تحتوى على استجابات حرة حقق النجاح فى اختبارات المهارات العملية ، الا ان هذه الوسائل لم تنجح نجاحا مؤكدا فى تصحيح الاختبارات اللغوية ، ومع ذلك فان التباين بين المصححين يمكن تقليله اذا أعددنا بروتوكولات توضح نمط التصحيح المتفق عليه بالنسبة للاجابات النموذجية ٠ ومن أمثلة ذلك نماذج التصحيح الخاصة باختبار ستانفرد ـ بينيه للذكاء ، وبروتوكولات اختبار رورشاخ واختبار تفهم الموضوع ، وقواعد التصحيح فى اختبارات الابتكار٠

وعلى الرغم من أن بروتوكولات التصحيح قد تكون مناسبة لأغراض القياس ، الا أننا نؤكد بعض الحذر حول تصحيح اختبارات الاستجابات الحرة بوجه عام ٠ فمعظم مصادر تباين الخطأ فيها تنشأ مما يمكن تسميته « المعادلة الشخصية » للمختبرين أو المصححين ، فقد لوحظ مثلا أن بعضهم يعطى درجات مرتفعة بصفة مستمرة بينما يميل البعض الآخر الى اعطاء درجات منخفضة ٠ كما أن الدرجة التى يعطيها المختبر او المصحح تتأثر باستعداده العقلى وتحيزه وتخمينه واغراضه الشخصية وأهوائه ، وفى هذه الحالة لا مناص من أن يقوموا لتصحيح اكثر من مصحح واحد تتوافر فيهم جميعا خصائص التدريب الجيد على بروتوكولات تصحيح هذه الاختبارات أو تقدم

لهم قواعد التصحيح واضحة بحيث يمكنهم تطبيقها دون الوقوع فى اخطاء سوء التأويل ٠

ولعل من المصادر الاخرى لتباين الخطأ فى هذا النوع من الاختبارات ما يسمى اثر او وضع الهالة الذى يحدثه المفحوص فى الفاحص ويمكن التقليل من هذا الأثر اذا لجأنا الى ما يسمى التصحيح « الأعمى » ، أى يقوم المختبر او المصحح بتقدير الدرجات للافراد دون ان يعلم نتائج الاختبارات السابقة التى اجريت عليهم أو بالفئة التى ينتمى اليها المفحوص فى البحث (المجموعة التجريبية أو الضابطة) أو المستوى الذى هو عليه (العقوق او التخلف) وبدون هذا الاحتياط يصعب علينا أن نتفادى تأثير هذا العامل الذاتى فى الدرجة المعطاة ٠

تصحيح الأسئلة التى تتطلب الاختيار من متعدد :

يتكون نموذج التصحيح لما نسميه اختبارات التعرف من قائمة الاجابات الصحيحة (أو مفتاح التصحيح) يمكن أن يستخدمه الفاحص بل تكاد تقول أن أى شخص يمكنه أن يستخدمه فى تصحيح هذه الاختبارات ، وقد اخترعت عدة وسائل اكثر كفاءة فى تسجيل الاستجابات فى هذا النوع من الاختبارات وتصحيحها ، ومنها النسخ الكربونية ومنها نجد ان الوجه الخلفى لصفحة الأسئلة (أو ورقة الاجابة) تلتصق به وتخفيه ورقة كربونية ، ويكون مفتاح الاجابة مطبوعا على هذا الوجه الخلفى ٠ ويختار المفحوص الاجابة ثم يضع العلامة التى تدل على هذا الاختبار على الوجه الظاهر لورقة الأسئلة (أو الاجابة) ، وبالطبع فان هذه العلامة تطبعها ورقة الكربون على الوجه الخلفى للورقة ، وعند التصحيح يقوم المختبر أو المصحح بفض الوجه الخلفى الذى توجد فيه مربعات مطبوعة توضح اين يجب أن تكون الاجابات الصحيحة (أى مفتاح الاجابة) ، وما على المصحح فى هذه الحالة الا عد العلامات الكربونية التى طبعت فى المربعات الصحيحة ٠ وبالطبع فان هذا الاجراء مكلف اقتصاديا ٠

ومن الاجراءات الشائعة استخدام أوراق مستقلة للاجابة ، وهو اجراء يؤدى الى اختصار التكاليف ، لأن كتيب الاختبار يمكن استخدامه فى هذه

الحالة اكثر من مرة ، كما أن الاجابات يمكن تصحيحها بسهولة بمفتاح مثقوب او بواسطة الآلات الحديثة ٠

وقد ازداد في الوقت الحاضر الاعتماد على تصحيح الاختبـــارات تصحيحا أوتوماتيكيا الى حد أن برامج كاملة للتقـــويم النفسى فى أوربا والولايات المتحدة تتم بالحسابات الالكترونية ٠ وتختلف طرق التصحيح الالى الا أنها تتشابه في عدد من العناصر منها ضرورة وجـــود ورقة اجـــابة بمواصفات معينة ، وأن يقوم المفحوص بتسويد مكان الاجابة بالقلم الرصاص، وتقوم الآلة بقراءة الاجابة بواسطة نظام للشفرة الموضعية باحدى طريقتين أولاهما خلايا ضوئية حساسة للاســود والابيض ٠ وثانيهما التوصـــيل الكهربائى لمادة الجرافيت (التى يصنع منها الاقلام) ، وفى الطريقة الأخيرة نجد الآلة تحتوى على « أصابع » مكهربة توضح علامات القلم الرصاص كما تحتوى الآلة على عداد تسجيل العدد الكلى للعلامات الموضوعية فى أماكنها الصحيحة ، كما أن الآلة يمكنها أن تسجل عدد الأخطاء ، كما يمكنهـــا أن تصحح الدرجة من أثر التخمين ، ويمكن لبعض هذه الآلات أن يصحح عدة الاف من أوراق الاجابة تصحيحا دقيقا فى الساعة الواحدة ٠ والحاسب الالكترونى يمكنه أن يقرأ ورقة الاجابة مرة واحدة ثم يقارن اجابات المفحوص بالمفتاح المختزن فى ذاكرة الكومبيوتر ، وتطبع درجة المفحوص اما على ورقة الاجابة ذاتها أو فى بطاقة خاصة ، كما يمكن للحاسب الالكترونى أن يقدم درجات المفحوصين مبوبة (مصنفة على النحو الذى يرغبه الفاحص) ، كما يمكنه أن يقوم بتحليل هذه الدرجات مباشرة بالطرق الاحصائية الملائمة٠

والصعوبة الرئيسية التى تواجهها آلات تصحيح الاختبارات ، والتى يهتم بها المختصون فى تطبيق الاختبارات ، هى أن هذه الآلات لا يمكنها أن تصحح تصحيحا دقيقا ما لم تكن الأماكن المخصصة للاجابة قد قام المفحوص بتسويدها تسويدا كاملا ونظيفا وحسب المواصـفات المطلوبة والا يكون قـد وضع اجاباته فى غير مواضعها ٠ فمن المعروف أن الآلة تعطى وحدة الدرجة عندما تظهر العلامة التى وضعها المفحوص فى المكان الصحيح ومع ذلك فنحن نتوقع أن تكون العلامة التى يضعها المفحوص بالقلم الخفيف أو الثقيل، أو تكون عريضة أو ضيقة ٠ ولذلك فلا بد أن يضع المفحوص هذه العلامة بقلم معين وأن يملأ المساحة الكاملة المخصصة للاجابة فى ورقة الاجابة حتى نتأكد أن

الآلة سوف تعدها ولا تتجاهلها ٠ ومن ناحية أخرى فان الآلة قد لا تتجاهل العلامات غير الكاملة والمحو غير النظيف للأخطاء ٠ وقد يكون عدد الاجابات فى ورقة الاجابة التى لا تتوفر فيها المواصفات اللازمة كثيرا ، ولذلك نجد فى بعض المؤسسات موظفين يقومون بفحص كل ورقة قبل أن تسلم للآلة لتصحيحها ، وهؤلاء يقومون بمراجعة اجابات المفحوصين لا بغرض تصحيحها ولكن بغرض التأكد من أن طريقة الاجابة مناسبة للتصحيح ، ومعنى ذلك زيادة تكاليف التصحيح ٠ ولا شك انه يمكن التغلب على هذه الصعوبة باجراء الاختبار وتطبيقه على نحو سليم ٠

تصحيح اثر التخمين :

من المشكلات الهامة فى تصحيح الاختبارات التى تتطلب اختيار الاجابة من بديلين او اكثر مايتصل بالتخمين guessing ، وهو مصطلح عام يشير الى انماط السلوك التى تصدر عن المفحوص حين يستجيب لأحد البدائل الاختيارية للاجابة على السؤال الذى لا يعرف اجابته الصحيحة ٠ وهذه الأنماط السلوكية عديدة ، تشمل من بين ما تشمل ما يلى (٣١ : ٥٩) ٠

١ ـ حذف بديل أو أكثر من بين البدائل الاختيارية للاجابة على السؤال والحكم عليها بأنها خاطئة تماما ، أو حين يكون المفحوص غير متأكد من صوابها ٠

٢ ـ الاستفادة من بعض المنبهات غير المقصودة والمتاحة بسبب صياغة السؤال أو البدائل الاختيارية ، وقد تكون هذه المنبهات فى المعنى او فى للبنية النحوية ٠

٣ ـ وقوع المفحوص فى الشرك الذى ينصبه له مؤلف الاختبار حين يطرح بديلا جذابا ولكنه خاطىء ٠

٤ ـ الاستجابة على اساس ما يوجد فى أحد البدائل الاختيارية من عناصر تجذبه دون أن يكون لديه ثقة كبيرة فى صحته ٠

٥ ـ استخدام طريقة عشوائية فى الاستجابة على نحو يعتمد كلية على

المصادفة ، وقد يصل ذلك فى تطرفه الى حد انتقاء البدائل الاختيارية التى تقع فى الاختبار فى موضع معين : مثل انتقاء الاجابة (أ) أو (حـ) أو (هـ) فى جميع الأسئلة •

ومن المؤكد أن المفحوص اذا لجأ فى أجابته الى محض التخمين معتمدا فى ذلك على العشوائية الكاملة فانه قد يحصل على بعض الدرجات ، لأن بعض تخميناته قد تقع بالمصادفة على الاجابات الصحيحة ، ولعل هذه المسألة تزداد وضوحا اذا علمنا انه فى بناء الاختبار الموضوعى الجيد الذى تقوم أسئلته على الاختيار من متعدد عادة ما تتوزع الاجابات الصحيحة على البدائل المختلفة بالتساوى ، فمثلا اذا كانت أسئلة الاختبار من نوع البديلين فان ٥٠٪ من الاجابات الصحيحة يكون رمزها (أ) مثلا ، ٥٠٪ من هذه الاجابات يكون رمزها (ب) ، على أن تتوزع عشوائيا ، دون أى تتابع منظم فى الاختبار كله • فاذا كانت أسئلة الاختبار من نوع البدائل الثلاثة فان ٣٣٪ من الأسئلة يكون اجاباتها الصحيحة (أ) ، ٣٣٪ رمزها (ب) ٣٣٪ رمزها (حـ) على أن تتوزع مرة أخرى عشوائيا فى الاختبار ، وبالمثل اذا كانت البدائل أربعة تتوزع الاجابات الصحيحة لكل ٢٥٪ من الأسئلة على أحد الرموز (أ) ، (ب) ، (حـ) ، (د) ، وهكذا •

لوافترضنا أن مفحوصا اعتمد على العشوائية الكاملة فاختار الاجابات التى رمزها (أ) فى جميع الأسئلة ، انه فى هذه الحالة يحصل دون حق على الدرجة ٥٠ فى الحالة الأولى ، ٣٣ فى الحالة الثانية ، ٢٥ فى الحالة الثالثة •

وحيث أننا لا نعرف فى هذه الحالة عدد التخمينات الصحيحة التى توصل اليها المفحوص بهذه الطريقة ، فاننا نستطيع استنتاج هذا العدد من عدد التخمينات الخاطئة والتى تمثلها المعادلة الآتية :

$$\text{خ خ} = \frac{\text{خ}}{\text{ن} - 1}$$

حيث خ خ = عدد التخمينات الخاطئة
خ = عدد الاجابات الخاطئة فى الاختبار كله
ن = عدد البدائل الاختيارية

ويمكن استخدام المعادلة السابقة فى معادلة اشمل لتصحيح اثر التخمين والتى تستخدم فى تقدير عدد الأسئلة التى يعرف المفحوص اجاباتها الصحيحة بالفعل ويستحق عليها الدرجة الاختبارية وهذه المعادلة هى :

$$د = ص - \left(\frac{خ}{ن - ١} \right)$$

حيث :

ص = عدد الاجابات الصحيحة فى الاختبار كله
د = الدرجة المصححة من اثر التخمين
خ = عدد الاجابات الخاطئة فى الاختبار كله
ن = عدد البدائل الإختيارية

تفسير الدرجات

الدرجات الخام :

تعطينا معظم الاختبارات وصفا كميا مباشرا لأداء الشخص ، ويطلق على الوصف الكمى للاداء اســـم الدرجة الخـــام raw score وقد تكون الدرجة الخام عدد الأسئلة التى اجاب عليها المفحوص فى اجابة صحيحة أو خاطئة أو الزمن الذى استغرقه فى الأداء او النســبة المئوية لعدد الاجابات الصحيحة او الخاطئة على العدد الكلى لأسئلة الاختبار ٠ ويفسر معظم الأفراد الدرجة الخام ويستخدمونها دون معرفة لجوانب القصور فيها لأن هذه الدرجات تتميز بالسهولة فى الحصول عليها وبالفتنا لها من خبرتنا الطويلة بالامتحانات المدرسية ٠ ومع ذلك فاننا نستطيع القول أن الدرجة الخام فى الاختبار النفسى لا معنى لها فى ذاتهـــا وليست لها اية دلالة ، ولا يمكن أن تفسر الا باللجوء الى اساس آخر للمقارنة ، ولعلنـــا نعود بهذا الى التعريف الأساسى للاختبار النفسى ٠ ومن المهم أن نؤكد فى هذا الصدد أننا لا نستطيع تفسير الدرجات فى الاختبار النفسى كما نفعل بالقيم التى نحصل عليها من مقياس الظواهر الطبيعية (الطول مثلا) ٠ فمقاييس الظواهر الطبيعية هى من نوع مقاييس النسبة ، أى لها صفر مطلق وتتألف من وحدات متساوية ، ولذلك فان هذه المقاييس تسمح لنا بتفســـير مباشر للقيم التى نحصل عليها منها ٠ فنقول مثلا ان طول الطفل (س) نصف طول والده (ص) باستخدام درجات المتر مباشرة ٠ اما فى القياس النفسى فاننا لا نستطيع مثل هذا القول لأن المقاييس فى هذه الحالة لا تكون فى أحسن حالاتها الا من نوع مقاييس المسافة ٠ لنفرض مثلا أن الطفل (س) حصل على درجة مقدارها ١٠٪ من أسئلة اختبار ، فى معانى الكلمات فهل يعنى ذلك أن ذلك الطفل لا يعرف الا عشر الكلمات التى يجب أن يعرفها ؟ الاجابة على ذلك بالنفى لأن الفاحص قد لا يكون مهتما الا بالكلمات الصعبة غير المألوفة للطفل ولهذا نجد أن محصوله اللفظى قد يكون أكبر فى اختبار اقل صعوبة ٠ وحتى لو حصل الطفل على الدرجة صفر فى هذا الاختبار فان ذلك يعنى أن الطفل ليست لديه قدرة لغوية على الاطلاق ٠ فالفرق بين (س) الذى حصل على الدرجة صفر فى الاختبار والطفل ص الذى يحصل على النهاية القصوى أى يجيب على اسئلة الاختبار جميعا ، هو فرق فى القـــدرة على

معرفة عينة من الكلمات يتكون منها الاختبار اشتقت من بين الاف الكلمات ، فالدرجة التى تدل على ٦٥٪ صواب فى أحد الاختبارات قد تتســاوى مع الدرجة التى تساوى ٣٠٪ فى اختبار آخر أصعب ، ومع ٨٠٪ فى اختبار ثالث اسهل • فمستوى صعوبة الكلمات التى يتألف منها الاختبار تحـــدد الدرجة الخام للمفحوص •

معنى ذلك أن الفرق بين الأفراد فى الدرجات الخام للاختبارات النفسية لا يدل بالضرورة على وجود مسافات « حقيقية » بين الأفــراد لأن هــذه الاختبارات لا يتحقق فيها الشرطان الأساسيان لمقاييس المسافة وهما الصفر المطلق والوحدات المتساوية • فمثلا فى اختبار الذكاء قد يحصل التلميذ (أ) على الدرجة ٥٣ والتلميذ (ب) على الدرجة ٥٦ والتلميذ (حـ) على الدرجة ٥٩ ، أى تكون الفروق بين الدرجات الخام متســاوية ، ومع ذلك لا يدل هذا على أن الفرق بين (حـ) و (ب) يساوى تماما الفرق بين (ب) و (أ) ، ولا يمكن الاجابة على مثل هذا السؤال اجابة مؤكدة ، لأن الفروق بين الدرجات انما تعتمد على طبيعة الأسئلة (أو الوحدات) التى تكون الاختبار كما قلنا • ولهذا لابد من أن نلجأ الى بعض الأسس التى تتجاوز الدرجات الخام فى المقارنة بين المفحوصين وتفسير درجاتهم والحكم على ادائهم •

ويضيف فؤاد أبو حطب (٩) هذه الأسس التى يجب استخدامها فى المقارنة ثم التفسير والحكم على الدرجة الخام الى ثلاثة أنواع هى :المعايير norms ، والمحـــكات Criteria ، والمستويات Standards وفيما يلى تفصيل الأسس الثلاثة :

(اولا) المعايير

كانت « المعايير » أول ما ظهر فى ميدان القياس النفسى للتغلب على الصعوبات المتضمنة فى تفسير الدرجات الخام والحكم على اداء المفحوص وتعد المعايير أساسا لتفسير أداء المفحوصين والمقارنة بينهم فى ضوء أدائهم الفعلى ، وتتحدد فى ضوء الخصائص الواقعية لهذا الأداء ، وتعتمد فى جوهرها على الأداء الاختبارى لعينة ممثلة للاصل الاحصائى موضع الاهتمام

تسمى عينة التقنين Standardization sample وهكذا تتحدد المعايير تجريبيا بما تستطيع مجموعة ممثلة من الأفراد أداءه (عينة ممثلة) ، ثم نشير الى الدرجة الخام التى يحصل عليها المفحوص فى ضوء توزيع الدرجات التى تحصل عليها هذه العينة ليتحدد موقعه ومكانته فى هذا التوزيع : هل تتفق درجته مثلا مع الأداء المتوسط لعينة التقنين ؟ هل تقع فى مستوى أقل قليلا (أو أعلى قليلا من المتوسط ؟ هل تقع فى الطرف الأعلى من التوزيع ؟ وهكذا .

وحتى يمكن أن نحدد على وجه الدقة موضع الفرد بالنسبة لعينة التقنين فان الاختبار النفسى يتحول الى نوع من مقاييس الرتبة أو مقاييس المسافة وتتحول الدرجة الخام الى هذا النوع من القياس عن طريق البحث عن مقابلها من رتبة أو مسافة وتسمى هذه القيم المقابلة للدرجات الخام بالدرجات المشتقة derived scores . وهذه الدرجات المشتقة تفيد فى تحقيق غرضين : أولهما تحديد الوضع النسبى للفرد فى العينة المعيارية وتقويم أدائه فى ضوء أداء الآخرين المماثلين له أو المختلفين عنه) ، وثانيهما أنها تعطينا مقاييس قابلة للمقارنة سواء بين الأفراد أو بين الاختبارات .

وحتى عهد قريب كانت المعايير هى الأساس الوحيد الذى يتطلبه علماء القياس النفسى للمقارنة بين الدرجات الخام وتفسيرها ، حتى أنها كانت تعد شرطا جوهريا لجودة الاختبارات النفسية على اختلاف أنواعها وأغراض استخدامها . الا أنه مع التطور المعاصر فى النظرية الأساسية للاختبارات النفسية وظهور أنواع جديدة منها وأغراض متنوعة لها ، وخاصة ما يسمى الاختبارات المرجعة الى المحك توجه الباحثون الى اعتبار المعايير أساسا واحدا من بين عدة أسس لهذه المقارنة ، واقتصر دورها على نوع واحد فقط من هذه الاختبارات هو ما يسمى « الاختبارات المرجعة الى المعيار » ، ونعرض فيما يلى المعايير الشائعة الاستخدام فى هذا النوع من الاختبارات .

المعايير الارتقائية او النمائية :

تشير المعايير الارتقائية و النمائية developmental الى المعنى الذى يعطى للدرجة الخام ويدل على اقتراب أو ابتعاد نمو المفحوص عن

المسار الطبيعي لنمو العملية النفسية موضع القياس • وعلى الرغم من شيوع هذا النوع من المعايير (وخاصة في صورة معيار العمر العقلي بالنسبة لاختبارات الذكاء) الا أنه ليس دقيقا ولا يصلح للمعالجة الاحصائية المحكمة ويجب الا تتجاوز استخدامه حدود الأغراض الوصفية في مجال البحث او المواقف الكلينيكية •

١ - معايير العمر والفرق الدراسية : لقد كان العالم الفرنسي الفريد بينيه Binet أول من استخدم مفهوم العمر العقلي mental age في تعديله لمقياسه المشهور في الذكاء عام ١٩٠٨ • وفي مقاييس العمر كما في عينة التقنين ، وتتطابق الدرجة العمرية التي يحصل عليها الفرد مع أعلى مستوى عمرى يمكن أن يصل اليه بنجاح • فاذا استطاع طفل عمره الزمنى ١٠ سنوات أن يجيب بنجاح على أسئلة ١٢ سنة فان عمره العقلي يصبح ١٢ سنة ، ومعنى ذلك انه يسبق عمره الزمنى بعامين • لأن اداءه يساوى اداء الطفل المتوسط من عمر زمنى أكبر منه بعامين •

وقد لوحظ أن اداء المفحوص في الاختبارات التي تعتمد على معيار العمر العقلي يظهر قدرا من التشتت • وبعبارة أخرى فان الطفل قد يفشل في اداءبعض الأسئلة التي تقع في مستوى عمرى أدنى من عمره العقلي ، كما يمكنه الاجابة على بعض الأسئلة التي تقع في مستوى أعلى منه • ولهذا السبب نلجأ عادة الى حساب العمر القاعدي basal age ويقصد به أدنى مستوى عمرى يستطيع الطفل أن يجيب على جميع أسئلته والأسئلة السابقة عليه ، ثم تضاف اليه درجات جزئية - في صورة كسور من السنة او شهور - تدل على جميع الأسئلة التي أجاب عليها من مستوى أعلى من العمر القاعدي وبذلك يصبح العمر العقلي للطفل هو مجموع العمر القاعدي والأشهر الاضافية التي تدل على ما حصل عليه الطفل من مستويات عمرية أعلى منه حتى يصل الى سقف الاختبار والذي يدل على المستوى العمرى الذي يعجز الطفل عن الاجابة على جميع أسئلته والأسئلة اللاحقة عليه •

وقد نلجأ الى استخدام معايير العمر العقلي في الاختبارات التي لا تنقسم مفرداتها الى مستويات عمرية كما هو الحال في اختبار ستانفرد - بينيه • وفي هذه الحالة تكون الدرجة الخام هي عدد الاجابات الصحيحة

أو الخاطئة على جميع أسئلة الاختبار ـ أو الزمن الكلى الذى استغرقه المفحوص فى حله وتحسب متوسطات الدرجات الخام التى تحصل عليها المجموعات العمرية المختلفة فى عينة التقنين ، وتكون تلك المتوسطات هى معايير الأعمار العقلية للاختبار كما تستخدم معايير عمرية أخرى مثل العمر الطولى والعمر الوزنى وعمر التسنين والعمر الرسغى وعمر قبضة اليد وعمر القراءة وغير ذلك من الخصائص النفسية والجسمية ذات الصبغة الارتقائية • ويمكن بالطبع استخدام هذا الاجراء للحصول على معايير الفرق الدراسية فى حالة الاختبارات التحصيلية المقننة حيث تستخدم فئات الفرق الدراسية بدلا من فئات الأعمار •

ومن أهم مشكلات وحدة العمر العقلى ووحدة الفرقة الدراسية كمعيار اختبارى أنهما لا تتقدمان بمعدل ثابت فى مختلف مراحل العمر أو مراحل الدراسة ، فالنمو العقلى أسرع وبالتالى يكون مقداره أكبر فى الطفولة المبكرة منه فى الطفولة المتأخرة ، ويزداد بطئا وبالتالى يقل مقداره باقتراب الشخص من حد النضج ، كما أن النمو التحصيلى ليس متكافئا فى جميع الفرق الدراسية ومعنى ذلك أن مقدار النمو الذى تتضمنه هذه الوحدات ـ بلغة القياس النفسى ـ يجعلها غير متساوية من عمر لآخر ، وبالتالى يجعلها غير قابلة للمقارنة •

وقد تتضح هذه الفكرة اذا أعطينا مثالا من العمر الطولى ، اننا فى هذه الحالة نجد أن الفرق بالسنتيمترات بين العمر الطولى للطفل فى سن ٣ أو ٤ أكبر من الفرق بين العمر الطولى فى سن ١٠ ، ١١ ، ومعنى ذلك أن عاما واحدا من « التفوق » أو « التخلف » فى سن ٤ سنوات يمثل اختلافا عن المستوى المتوسط أكبر كثيرا من مقدار يساويه من التفوق أو التخلف فى سن ١١ سنة •

٢ ـ نسبة الذكاء التقليدية : ابتكر علماء النفس معيار نسبة الذكاء intelligence quotient أو I.Q. للحصول على تفسير موحد للذكاء بصرف النظر عن العمر الزمنى للمفحوص • وقد أشار الى أهمية ما يسمى السنة العقلية شترن وتولمان • الا أن هذا المعيار استخدم لأولا مرة فى

مقياس ستانفورد ـ بينيه الذى أعده ترمان عام ١٩١٦ ويمكن الحصول على نسبة الذكاء كما يلى :

$$\text{نسبة الذكاء} = \frac{\text{العمر العقلى}}{\text{العمر الزمنى}} \times ١٠٠$$

فاذا تعادل العمر العقلى والعمر الزمنى تكون نسبة الذكاء فى هذه الحالة ١٠٠ وتدل على اداء متوسط أو عادى ، واذا قلت عن ١٠٠ دلت على تخلف أو ضعف ، واذا زادت عن ١٠٠ دلت على التفوق ٠

وتفيد نسبة الذكاء فى ضبط الضمور العقلى لطفل عمره الزمنى ٤ سنوات هو ٣ تصبح نسبة ذكائه فى هذه الحالة ٧٥ فاذا بلغ هذا الطفل سن ١٢ وكان عمره العقلى ٩ تظل نسبة ذكائه ثابتة أى ٧٥ ٠ اى أن نسبة الذكاء تدل على نفس الوضع النسبى للفرد فى الجماعة سواء فى سن ٤ سنوات أو فى سن ١٢ سنة ٠ وبعبارة أخرى فان نسبة الذكاء يمكن استعمالها فى المقارنة فى مختلف الأعمار ، بمعنى أن تفسير نسبة الذكاء يظل هو نفسه بصرف النظر عن العمر الزمنى للمفحوص ٠

ولابد أن نشير فى هذا الصدد الى أنه لكى تظل نسبة الذكاء ثابتة فان ذلك لا يمكن ـ نظريا ـ الا فى حالة واحدة فقط وهى أن وحدة العمر العقلى تتناقص (أى تصبح أصغر) تناقصا يتناسب تناسبا مباشرا مع العمر الزمنى ٠ وهذا شرط ضرورى حتى يمكن استخدام نسبة الذكاء استخداما له معنى ٠

وحين تتناقص وحدة العمر العقلى تبعا للتقدم فى العمر على هذا النحو فان الفروق الفردية التى تقاس بهذه الوحدة يفترض فيها بالطبع أنها تتزايد بنفس النسبة ٠ ويشبه ذلك من الناحية الحسابية أن الفروق الفردية فى الطول تكون أكبر ١٢ مرة حينما تقاس بالبوصات عنها لو قيست بالأقدام وذلك لأن البوصة هى $\frac{١}{١٢}$ من القدم وبنفس المنطق فان الفروق الفردية أو التباين فى العمر العقلى يجب أن تكون فى سن ١٤ ضعفها سن ٧ سنوات، ولا يمكن لنسب الذكاء أن تظل ثابتة الا اذا تزايد الانحراف المعيارى للاعمار

العقلية بنفس نسبة زيادة الأعمار الزمنية ، وبهذا الشرط وحده يمكن أن يكون لنسبة الذكاء نفس المعنى فى مختلف الأعمار ٠

وقد أمكن تحقيق هذا الشرط فى الطبعات الحديثة من اختبار ستانفرد ـ بينيه فقد اتضح من دراسة انتشار الأعمار العقلية من سن ٦ حتى سن ١٨ أن هناك اتجاها نحو زيادة التباين فى الأعمار العقلية تبعا لزيادة الأعمار الزمنية ٠ ولكننا نجد فى كثير من الأحيان عدم تساوى الانحرافات المعيارية لنسب الذكاء فى مختلف الأعمار ، وهذه الصعوبة أمكن التغلب عليها فى اختبار ستانفرد ـ بينيه باعداد جدول تصحيح يستخدم فى تصحيح نسب الذكاء فى مستويات عمرية معينة ٠

وهناك مشكلة أخرى فى تطبيق معيار نسبة الذكاء نجدها فى اختبار البالغين والكبار فالواقع أن مفهوم معايير العمر على أطلاقه يدل دلالة مباشرة على أن فائدته مقصورة على الأطفال لما يتضمنه من معانى ارتقائية ٠ أما بالنسبة للراشدين والكبار فان ظواهر النمو ليست على نفس الدرجة من الوضوح الا فى حالات تدهور الشيخوخة ولذا نجد أن فى اختبار مثل ستانفرد ـ بينيه لا يتحسن أداء البالغ المتوسط كثيرا ٠ صحيح انه قد أضيف الى الاختبار مستويات رفيعة لذكاء الكبار حتى يتوافر فيه سقف مناسب وبالتالى يمكننا الحصول على أعمار عقلية أعلى ، ولكننا اذا وجدنا شخصا معينا يحصل على عمر عقلى مقداره ٢٠ فى هذا الاختبار فان ذلك لا يسمح لنا بتفسير مؤكد وعلى نفس الدرجة من الوضوح كما هو الحال فى الاعمار العقلية الدنيا ٠ فمن المؤكد أن العمر العقلى ٢٠ لا يمكن تحديده بما يمكن أن يؤديه الشخص المتوسط البالغ من العمر ٢٠ سنة ، فى هذا الاختبار ، أما اذا اختبرنا شخصا بالغا ضعيف العقل يقل عمره العقلى عن حدود النضج فان استخدام مفهوم العمر العقلى فى هذه الحالة عنه يتشابه باستخدامه مع الاطفال ٠ اما فى حالة الراشدين والكبار من الأسوياء أو المتفوقين فلا مناص من استخدام أنواع أخرى من المعايير ليس من بينها معيار العمر لصعوبة تحديدهم عندهم ٠

اما المشكلة الثالثة التى تواجهنا حينما نفسر الدرجات الخام بدرجات العمر تتمثل فى أن هذا النوع من المعايير لا يستخدم الا فى الوظائف النفسية

والعمليات العقلية التى تتغير تغييرا واضحا ومتسقا مع النمو ٠ اما بالنسبة للسمات التى لها علاقة ضئيلة بالعمر الزمنى او لا تكون لها علاقة مباشرة به لا يجوز بالطبع استخدام معيار العمر معها ، ومن أمثلة ذلك مقاييس سمات الشخصية ، فبعض هذه السمات يظل ثابتا فى معظم مراحل نمو الفرد ٠

ولأسباب تعود الى هذه المشكلات وغيرها توقف علماء القياس النفسى المعاصرون عن استخدام نسبة الذكاء على النحو المشار اليه ، وأصبح يطلق عليها اسم نسبة الذكاء التقليدية تمييزا لها عن انواع أخرى من المعايير اكبر دقة سنشير اليها فيما بعد ولعل فى هذا ما ينبه الباحثين فى مصر والبلاد العربية الى أن يتوجهوا فى بناء الاختبارات النفسية وفى استخداماتها فى الأغراض المختلفة (وخاصة الأغراض الكلينيكية) الى المعايير الآخرى المناسبة غير نسبة الذكاء التقليدية ٠

٣ ـ **المئينيات** : يعبر عن الدرجات المئينية percentiles فى صورة نسب مئوية للافراد من عينة التقنين الذين يقعون أدنى من درجة خام معينة ٠ وتدل على الوضع النسبى للفرد بمقارنته بعينة التقنين هذه ٠ وهى تعد من نوع الرتب فيما عدا أننا فى الترتيب نبدأ عادة باعتبار الشخص الذى يقع فى القمة ـ أى أفضل أفراد المجموعة من حيث الأداء ـ فى الرتبة الأولى، بينما فى المئينيات فاننا نبدأ العد من أسفل ، وبالتالى فان ادنى المئينيات يدل على أسوأ مكانة للفرد ٠

ويقابل المئينى الخمسون منتصف جماعة التقنين او وسيطها فاذا زاد المئينى عن ٥٠ يدل هذا على أداء أعلى من المتوسط ، واذا قل عنه دل على أداء أقل من المتوسط ٠ ويعرف المئينى ٢٥ ، والمئينى ٧٥ احصائيا بالربيع الأدنى والربيع الأعلى من التوزيع ، وهما مثل الوسيط يعطيانا معالم تصف توزيع الدرجات ومقارنته بغيره من التوزيعات ، وهى كالوسيط والربيع فى طريقة حسابها ، كما أنها تقسم التوزيع التكرارى الى نقاط ، وهى فى حالة المئينيات تصل الى ٩٩ نقطة ٠

ويجب الا نخلط بين المئينيات ودرجات النسب المئوية الشائعة ، فدرجات النسب المئوية هى نوع من الدرجات الخام يعبر عنها فى ضوء نسب مئوية

للاجابات الصحيحة • أما المئينيات فهى درجات مشتقة يعبر عنها فى ضوء نسب مئوية للمفحوصين فاذا افترضنا أن لدينا ٤٠ شخصا منهم ٢٧ حصلوا على درجات أعلى من الشخص (أ) ، ١٢ حصلوا على درجات أقل منه ، فاننا فى هذه الحالة نضع هذا الشخص بين المجموعتين معتبرين أنه تفوق على ٣١٪ من العدد الكلى للمجموعة ، وتكون درجته المئينية • والدرجة الخام التى تقل عن أى درجة حصل عليها افراد عينة التقنين تكون رتبتها المئينية صفرا ، والدرجة الخام التى تزيد على أعلى درجة حصل عليها افراد عينة التقنين تكون رتبتها المئينية ١٠٠ وهذه المئينيات لا تتضمن بحال من الأحوال أن تكون الدرجات الخام المقابلة لها مقدارها صفرا أو تدل على الأداء الكامل أو الحد الاقصى للدرجة الخام •

وللمئينيات ميزات عديدة • فهى سهلة الحساب ميسورة الفهم ، كما أن استخدامها يصلح لكل أنواع الاختبارات ، كما تصلح للاطفال والمراهقين والراشدين على حد سواء • الا أن أهم مشكلاتها تنشأ من عدم تساوى وحداتها وخاصة عند أطراف التوزيع ، فاذا كان توزيع الدرجات الخام يقترب من التوزيع الاعتدالى فان الفروق بين الدرجات الخام الأقرب الى الوسيط أو مركز التوزيع تتزايد (أى تنتشر الحالات الكثيرة التى تقع فى المنتصف وتتفتت الى مئينيات كثيرة) بينما تتضاءل هذه الفروق عند الأطراف (أى تتجمع الحالات القليلة التى تقع عند الأطراف بعضها الى بعض فى عدد قليل من المئينيات) • فاذا أعدنا رسم التوزيع التكرارى بيانيا باستخدام المئينيات نحصل على ما يسمى التوزيع المستطيل •

ولا بد أن نشير الى أننا لا نستطيع استخراج متوسطات الدرجات المئينية ، لأن هذا المتوسط يختلف فى مقداره عما لو استخرجنا المتوسط الحسابى للدرجات الخام ثم حولناه الى درجة مئينية • وبدلا من اللجوء الى المتوسط الحسابى لقياس النزعة المركزية للدرجات المئينية يمكن للباحث ان يعتمد على الوسيط باعتباره القياس الاحصائى المناسب للنزعة المركزية فى هذه الحالة

كما يجب الاشارة الى أن الدرجات المئينية لا يمكن المقارنة بينها ما لم تكن هذه المقارنة ممكنة من الجماعات التى تشتق منها فاذا اختلفت

الجماعات ، كأن تكون احداها مثلا من تلاميذ مدرسة ثانوية صناعية والثانية من عمال تلمذة صناعية والثالثة من طلاب كلية للهندسة فان الدرجات الخام التى تقابل مئينى معين فى كل حالة من هذه الحالات تختلف ، كما أن الدرجة الخام الواحدة فى كل مجموعة من هذه المجموعات الثلاث تقابل مئينيات مختلفة ومن ذلك مثلا أن الدرجة الخام التى تقابل الدرجة المئينية الـ ٦٠ مثلا لطلاب الجامعة قد تقابل مثلا المئينى الـ ٧٥ فى مجموعة طلاب المدرسة الثانوية الصناعية وتقابل المئينى الـ ٩٠ فى مجموعة عمال التلمذةالصناعية· بعبارة أخرى اذا لجأنا الى استخدام الدرجات المئينية لابد للباحث أن يضع فى اعتباره عينة التقنين التى تشتق منها هذه المعايير ، ولو أن ذلك لا يحول دون استخدام مجموعة من جداول التحويل الى درجات مئينية تسمح للمختبر أو الباحث أن يقارن الشخص الذى يختبره بمجموعات مرجعية مختلفة ·

٤ ـ الدرجات المعيارية : يتزايد استخدام الدرجات المعيارية Satndard scores كمعايير فى الاختبارات الحديثة ، وتعتبر هذه المعايير أفضل صورة لتحويل الدرجات الخام لأنها تعبر عن بعد الدرجة الخام التى يحصل عليها المفحوص عن المتوسط منسوبا الى الانحراف المعيارى للتوزيع، أى أن :

$$\text{الدرجة المعيارية} = \frac{\text{الدرجة الخام} - \text{المتوسط}}{\text{الانحراف المعيارى}}$$

$$\text{أو } ذ = \frac{س - م}{ع}$$

ويطلق على هذا النوع «الدرجات المعيارية الخطية Linear standard scores ويقصد بذلك أننا نحتفظ بالعلاقات العددية الدقيقة للدرجات الخام الأصلية ، فنحن نطرح من الدرجة مقدارا ثابتا (المتوسط) ، ونقسمها على مقدار ثابت (الانحراف المعيارى) · ومعنى ذلك أن السعة النسبية للفروق بين الدرجات

المعيارية التى تشتق من هذا التحويل الخطى تتطابق تماما مع الفروق بين الدرجات الخام ، كما تتكرر جميع خصائص التوزيع الأصلى للدرجات الخام فى توزيع هذه الدرجات المعيارية ٠

ولعلنا نذكر أن من أهم أسباب تحويل الدرجات الخام الى درجات مشتقة من المعايير عامة هو جعل الدرجات التى نحصـــل عليها من مختلف الاختبارات قابلة للمقارنة ، وهذا الفرض لا يمكن أن يتحقق للدرجات المعيارية الخطية الا اذا كانت التوزيعات التكرارية للاختبارات المختلفة متشابهة ، فاذا كان أحد التوزيعين اعتداليا والاخر ملتويا فلابد فى هذه الحالة من اجراء تحويلات غير خطية للدرجات الخام حتى تتواءم مع نمط معين من أنماط التوزيع ، وعادة ما يختار الباحثون فى مجال الاختبارات المرجعة الى المعيار التوزيع الاعتدالى لأسباب نظرية خالصة ، ومعنى ذلك أن التوزيعات المختلفة يجب أن تتحول فى هذه الحالة الى توزيعات تتفق مع خصـائص المنحنى الاعتدالى ، ثم تحول الدرجات الى درجات معيارية اعتدالية باســتخدام جداول مساحات وارتفاعات المنحنى الاعتدالى normalized standard scores وتصبح لهذه الدرجات المعيارية الاعتدالية فى هذه الحالة نفس خصائص الدرجات المعيارية الخطية (أى متوسطها صفر وانحرافها المعيارى واحد صحيح) ٠

الا أن ما يجب أن ننبه اليه هو أن الدرجات المعيارية الخطية لا تتطابق مع الدرجات المعيارية الاعتدالية الا فى حالة واحـــدة فقط وهى أن تكون التوزيعات التكرارية التى حسبت منها الدرجات المعيارية الخطية اعتدالية بالطبع ٠ كما ننبه أيضا الى أن التحويلات الاعتدالية لا يجب أن تتم الا اذا توافرت شروط الاعتدالية من حيث السمة المقيسة والعينة المختبرة والمقياس المستخدم ، وهى شروط فصلها فؤاد أبو حطب فى موضع آخر (٩) ٠

ومن الواضح أن حساب الدرجات المعيارية الخطية يتضمن أن تكون هذه الدرجات سالبة أو موجبة ٠ فالدرجات السالبة تدل على أداء أدنى من المتوسط ، والدرجات الموجبة تدل على أداء أعلى منه ، والدرجة الخام التى تساوى المتوسط تقابل الدرجة المعيارية صفرا ٠ كما أننا نلاحظ على هذه

الدرجات المعيارية أنها تتضمن كسورا عشرية حتى تتوافر لدينا معلومات كافية نميز بها بين مختلف المفحوصين وخاصة اذا علمنا أن المدى الكلى لأغلب المجموعات لا يتعدى نطاق ٣ انحرافات معيارية ٠

وللتغلب على هاتين الصعوبتين اقترحت عدة تعديلات تهدف الى تحويل الدرجات المعيارية الى صورة أكثر ملاءمة وأيسر فى الفهم ٠ وتعتمد هذه التعديلات على بعض الخصائص الرياضية للدرجات المعيارية ٠ والتى تتمثل فى أن متوسط هذه الدرجات يساوى صفرا ، وانحرافها المعيارى يساوى الواحد الصحيح ، ولذلك نستطيع أن ندخل عليها ما نشاء من تعديلات باستخدام متوسطات وانحرافات معيارية جديدة ، وفى تحويل الدرجات المعيارية الأصلية الى المقياس الجديد يتطلب الأمر ببساطة أن نضربها فى الانحراف المعيارى الجديد ونجمعها الى المتوسط الجديد ٠ أى أن الدرجة المعيارية المعدلة تحسب كما يلى :

الدرجة المعيارية المعدلة = الدرجة المعيارية الأصلية × الانحراف المعيارى الجديد + المتوسط الجديد ٠

ذَ = (ذ × عَ) + مَ

وقد كثرت الدرجات المعيارية المعدلة وتعددت ، ويلخص الجدول رقم (٦) بعض أنواعها الشائعة ، وقد يكون أكثرها شيوعا فى الوقت الحاضر ، وبخاصة فى اختبارات الذكاء ، نسبة الذكاء الانحرافية ، وهى تختلف تماما عن نسبة الذكاء التقليدية التى عرضناها آنفا ، فنسبة الذكاء الانحرافية درجة معيارية معدلة متوسطها الجديد ١٠٠ وانحرافها المعيارى الجديد قد يكون ١٥ (كما هو الحال فى الطبعات الحديثة من مقياس ستانفرد ـ بينيه) أو ١٦ (كما هو الحال فى اختبارات وكسلر) ٠

وقد أثر بناة اختبارات الذكاء بمفهوم نسبة الذكاء فى معيارهم القائم على فكرة الدرجة المعيارية (نسبة الذكاء الانحرافية) لأسباب تاريخية ترجع فى جوهرها لاستخدام المصطلح لفترة طويلة فى علم النفس ، واختير الانحراف المعيارى الجديد (١٦ أو ١٥) فى ضوء التحليل الاحصائى

للانحرافات المعيارية لنسب الذكاء التقليدية والتى بلغ وسيطها فى مقياس ستانفرد - بينيه طبعة ١٩٢٧ حوالى ١٦ ، أما المتوسط ١٠٠ فقد اختير لأنه يشير فى نسبة الذكاء التقليدية الى المتوسط ٠ وهكذا تتوافر فى نسبة الذكاء الانحرافية (كدرجة معيارية معدلة) جميع الخصائص « الشكلية » لنسبة الذكاء التقليدية الا انها تختلف عنها تماما فى الأسس الرياضـــية وطرق الحساب كما بينا بالاضافة الى اختلافهما فى المعنى التفسيرى فنسبة الذكاء التقليدية هىنسبة العمر العقلى الى العمر الزمنى، أما نسبة الذكاء الانحرافية فتدل على اقتراب أو ابتعاد عن المتوسط بمسافات من الانحراف المعيارى ٠

جدول رقم (٦) بعض أنواع الدرجات المعيارية المعدلة

المتوسط	الانحراف المعيارى	الدرجةالمعياريةالتى تقابلانحرافا معياريا واحداأعلى من المتوسط	الدرجةالمعياريةالتى تقابلانحرافين معيارين أقل من المتوسط	اسم الدرجة المعيارية
صفر	١	١	- ٢	الدرجة المعيارية الأصلية
٥	٢	٧	١	الدرجة التســاعية أو التساعى المعيارى
٥٠	١٠	٦٠	٣٠	الدرجة التائية
١٠٠	١١٥و١٦	١١٥ أو ١١٦	٧٠ أو ٦٨	نسبة الذكاء الانحرافية

تعليق على المعايير : المعايير اذن هى أسس للحكم على أداء المفحوصين والمقارنة بينهم فى ضوء أدائهم الفعلى فى الاختبارات ولأنها ذات طبيعة كمية يمكن استخدام الحاسبات الالكترونية معها بحيث يقدم لنا الكومبيوتر تفسيرا للدرجات الخام ٠ وحتى يمكن المقــارنة بين المعايير الخاصــة بالاختبارات المختلفة ، او بين المعيار الواحد المستخدم مـع المفحـوص فى مراحل مختلفة من تطبيق الاختبار ، لابد من أن يكون الاختبار المســتخدم والذى تشتق من درجات عينة تقنيه المعايير الخاصة به معلوما للفاحص ٠ فالمعايير لا تستخدم لأغراض المقارنة المجردة المطلقة ٠

وتفيدنا معرفة طبيعة الاختبار المستخدم فى القياس فى تحديد محتواه

وما يقيسه ، ونود أن ننبه هنا الى الاختبارات التى تبدو لنا متشابهة فيما تقيس (كاختبارات الذكاء) هى فى الواقع مقاييس لعمليات نفسية مختلفة ٠

وحتى يدرك الفاحص مغزى المعايير ومعناها عليه أيضا أن يتحقق من القيم العددية المستخدمة فى حساب المعايير ٠ فحتى يستطيع المقارنة بين نسب الذكاء الانحرافية لاختبارين للذكاء لابد أن تكون واضحة تماما وحدات القياس المستخدمة فى حساب هذا المعيار ـ فنسب الذكاء الانحرافية (المتوسط جديد مقداره ١٠٠) تختلف تبعا للقيم المختارة للانحراف المعيارى الجديد (١٥ أو ١٦ مثلا) ٠

ولكى يفهم الفاحص طبيعة المعيار ويستخدمه استخداما صحيحا لابد أن يتعرف بعناية على عينة التقنية المستخدمة فى حساب معايير الاختبار ، وعلينا دائما أن نعلم أن معايير أى اختبار لا يمكن أن تتجاوز حدود خصائص عينة تقنينه بشرط أن تتوافر فيها دائما شروط العينة الجديدة وخاصـــة شرط التمثيل ، فاذا أردنا استخدامه مع مفحوصين ذوى خصائص مختلفة لابد من اعادة تقنينه على عينات جديدة ٠وهنا تجدنا ازاء موقف هام بالنسبة للمعايير ، وهو أن الاختبار الواحد قد تكون له معايير متعددة بعضها قد يكون عاما على المستوى القومى أو أقل عمومية على المستوى المحلى ، بل قد يصل الأمر الى حساب معايير خاصة لعينات تقنين مختلفة تشـــتق من مختلف الفئات التى يصنف اليها المجتمع الواحد أو الثقافة الواحدة ، وفى هذه الحالة ينسب المفحوص الى الأصل الاحصائى الذى ينتسب اليه ويكون الحكم على أدائه وتفسيره أقرب الى العدالة ٠

صحيح أنه ظهرت فى السنوات الأخيرة بعض الاتجاهات العمليـــة والنظرية تعين على بناء مقاييس موحدة « متحررة من أثر العينة » وتصلح للتطبيق على الأفراد والجماعات من مستويات مختلفة فى القدرة ولقياس محتوى من مستويات مختلفة من الصعوبة ، ومن هذه الطرق ما يســمى استخدام المجموعة المرجعية الثابتة fixed reference group ونمـــاذج السمة الكامنة latent trait models ، وهى تعتمد على مجموعة من الأسئلة أو الاختبارات الفرعية تسمى أسئلة أو اختبارات الارتكاز anchor

وتختار بموصفات معينة بحيث تتجاوز العينات والاختبارات والأسئلة ، الا أن هذه الطرق جميعا تهتم باحد مصادر الفروق بين الأفراد والجماعات وهو صعوبة الأسئلة ، وتوجد مصادر أخرى لا تقل أهمية ٠

وتبقى أخيرا الاشارة الى أن المعايير كأسس للحـــكم على الدرجات الخام وتفسيرها تصلح للاختبارات التى تسمى فى الوقت الحاضر الاختبارات المرجعة الى المعيار دون سواها ، ولا تصلح للانواع الاخرى التى تستخدم معها المحكات والمستويات كما سنبين فيما بعد ٠

المحكات :

يمكن للدرجات الخام أن تفسر أيضا فى ضوء الأداء فى محك متوقع، والمحك كما أشرنا من قبل هو أساس خارجى مستقل للحكم على الأداء فى الاختبار ، كما هو الحال فى برامج التدريب أو مقاييس الكفـــاية التعليمية والمهنية ، ويتفق هذا المعنى للمحك مع استخدامه فى بحوث الصدق ، والى هذا المعنى يجب أن يشير ما يسمى فى الوقت الحاضر الاختبارات المرجعة الى المحك الا أن المعنى الاصطلاحى الشائع لهذا النوع من الاختبـــارات يخرجها من المعنى الحالى للمحك ويدخلها فى فئة جديدة سنشير اليها تحت اسم « المستوى » ٠ وعلى ذلك فان أسس الحكم على ما يسمى « الاختبارات المرجعة الى المحك » ليست المحكات كما شاعت فى مجال القياس النفسى ، وانما هى المستويات ، وكان الاجدى لها أن تسمى « الأختبارات المرجعة الى المستويات » حتى لا توقعنا فى هذا الخلط السيمانتى الذى لا مبرر له ٠ وعلى كل فقد آثرنا الاشارة اليها بالاصطلاح الذى شاع فى الوقت الحاضر على أن يكون القارىء واعيا طول الوقت فى أن ما يقصد بالمحك فى هذا النوع من الاختبارات هو المستوى ٠

أما المحكات حين تستخدم فى الحكم على الأداء الاختبارى وتفسيره ، فان ذلك توسيعا لمعناها فى بحوث الصدق وامتدادا له الى عالم الواقع ، ولعل ما يتفق مع هذا المعنى ما اقترحه فؤاد أبو حطب من اشارة الى الصدق

المرتبط بالمحكات بالصدق العملى التطبيقى (*) • ويتضح ذلك من عرض اشهر الطرق التى تستخدم فيها المحكات فى الحكم على الدرجة الخام فى الاختبارات النفسية وتفسيرها وهى ما يسـمى طريقة جـداول التوقع Expectancy tables .

جداول التوقع : سبق أن قلنا أن كلا من الصدق التنبؤى والصـدق التلازمى يعتمد على تقويم درجات الاختبار فى ضوء محك مسـتقل يهدف الاختبار الى التنبؤ به أو تشخيصه على التوالى • والعلاقة بين درجات المفحوصين فى الاختبار ومكانتهم فى المحك يمكن أن يحللها الباحثون بطرق مختلفة ، وابسط صورة للتعبير عن هذه العلاقة هى ما يسمى جدول التوقع ومثل هذا الجدول يوضح توقع احتمال حصول الفرد على درجة معينة فى المحك (وليكن النجاح المدرسى أو المهنى) من حصوله على درجة معينة فى الاختبار •

وقد يأخذ جدول التوقع صورة رسم بيانى للمحك ذى البعدين (نجاح او فشل مثلا) يوضح العلاقة بين النسب المئوية للاشخاص الذين استبعدوا من التدريب المبدئى على الطيران وبين الدرجة التساعية التى يحصل عليها الشخص فى بطارية انتقاء الطياريين التى اعدها السلاح الجوى الأمريكى • ويلاحظ من هذا الجدول أن ٧٧٪ من الأشخاص الذين حصلوا على التساعى الأول استبعدوا بينما نجد أن حوالى ٤٪ فقط من اولئك الذين حصلوا على التساعى التاسع فشلوا فى اتمام التدريب • كما يتضح لنا من الجدول أن النسب المئوية تتزايد تدريجيا فى التساعيات التى تقع بين الأول والأخير •

والنسب المئوية المعطاة فى مثل هذا الرسم تعطى تقـديرا لاحتمال حصول الأفراد الذين تختبرهم على مكانة معينة فى المحك • فنحن نتنبأ مثلا من هذا الشكل أن حوالى ٤٠٪ من طلاب الطيران الذين يحصلون على الدرجة التساعية ٤ سوف يفشلون وأن حوالى ٦٠٪ منهم سـوف ينهون برنامج

(*) استخدم فؤاد ابو حطب هذا المصطلح فى سلسة محاضراته التى القاها عامى ١٩٧٩ _ ١٩٨٠ ، ١٩٨٠ _ ١٩٨١ بالمركز القومى للبحوث الاجتماعية والجنائية بالقاهرة عن الجوانب السيكومترية للاختبارات الاسقاطية •

التدريب على أعمال الطيران بنجاح • ويمكن تقرير توقعات مشابهة بنجاح أو فشل الأفراد الذين يقعون فى أى من المعايير التساعية • وعلى ذلك فالفرد الذى يحصل على درجة تساعية مقدارها ٤ تكون فرصته $\frac{٦}{٤}$ أى $\frac{٣}{٢}$ فى انهاء برنامج التدريب على أعمال الطيران بنجاح •

وعندما يكون المحك متصلا وليس من النوع ذى البعدين فحسب فان جدول التوقع يمكن اعداده اعدادا مباشرا من الرسم البيانى للانتشار ، وفى اعداد رسم الانتشار يحدد لكل فرد درجتاه فى الاختبار وفى المحك وفيه يمثل الخط الأفقى مقاييس المحك (تقديرات التلاميذ فى مادة البلاغة) ويمثل الخط الرأسى الدرجات فى اختبار الجمل • وفى هذا الاختبـــار يطلب من المفحوص أن يحدد الأخطاء فى النحو أو علامات التوقيف أو اســتعمال الكلمات فى سلاسل من الجمل • وكل علامة فى الشـــكل تدل على كل من الدرجة فى الاختبار والتقدير فى المحك لكل طالب من مائة ، كما أن الرسم يتضمن العدد الكلى لتكرارات فى الخلايا المختلفة وكذلك المجموع الخام للتكرارات •

ولتحويل جدول الانتشار هذا الى جدول توقع تحول التكرار الموجود فى كل خلية الى نسبة مئوية من المجموع الخام المقابل له • فمثلا يوجد ١٦ شخصا حصلوا على الدرجات بين ٣٠ و ٣٩ فى الاختبار ، من بينهم ٦٪ (شخص واحد) حصل على التقدير (هـ) فى مقرر البلاغة و ١٩٪ (٣ أشخاص) حصلوا على التقدير (د) و ٥٦٪ (٩ حالات) حصلوا على التقدير (حـ) و ١٩٪ (٣ أشخاص) حصلوا على التقدير (ب) • وبنفس الطريقة يمكن أن تحول جميع تكرارات الخلايا فنحصـــــل على جـــدول التوقع الآتى :

جدول (٧)

جدول توقع يوضح العلاقة بين الدرجات فى اختبار الجمل
وبين تقديرات مادة البلاغة

درجات الاختبار	النسب المئوية للافراد الذين حصلوا على كل تقدير				
	هـ	د	حـ	ب	أ
٨٠ ـ ٨٩					١٠٠
٧٠ ـ ٧٩				٢٠	٨٠
٦٠ ـ ٦٩			١٤	٦٣	٢٣
٥٠ ـ ٥٩			٣٩	٣٥	٢٦
٤٠ ـ ٤٩		١٤	٥٩	٢٧	
٣٠ ـ ٣٩	٦	١٩	٥٦	١٩	
٢٠ ـ ٢٩	١٣	٥٠	٣٧		
١٠ ـ ١٩		١٠٠			
صفر ـ ٩		١٠٠			

ومن هذا الجدول يمكن التنبؤ بالتقدير الذى يحصل عليه الفرد فى المحك من الدرجة التى يحصل عليها فى الاختبار . فمثلا يمكن أن نتوقع أن الشخص الذى يحصل على درجة أعلى من ٨٠ أن يحصل على تقدير (أ) والذى يحصل على درجة أعلى من ٧٠ اما أن يحصل على تقدير (أ) أو (ب) ، كما نتوقع أن الشخص الذى يحصل على درجة أعلى من ٥٠ لن يحصل على تقدير (هـ) أو (د) . وهناك عديد من التوقعات التى يمكن الوصول اليها من هذا الجدول .

ولا بد أن نشير فى هذا الصدد أن مثل هذه التنبؤات تتعرض لأخطاء العينة وخاصة اذا كان عدد الأفراد الذين نختبرهم صغيرا . وحيث أن النسب المئوية الفردية تعتمد على عدد صغير نسبيا من الأفراد فى كل خلية فان ذلك يعنى أن التغيرات الناتجة عن الصدفة عند الانتقال من عينة لأخـرى قد تكون كبيرة .

المســتويات :

النوع الثالث من الأسس التى تستخدم فى الحكم على الأداء الاختبارى

وتفسيره ما يسميه فؤاد أبو حطب (٩) المســـــتويات standards وهى ـ كما ذكرنا آنفا تتشابه مع المعايير فى انها أسس للحكم على الأداء فى ضوء هذا الأداء ذاته وليس فى ضوء أسس خارجية عنه ، الا انها تختلف عن المعايير فى انها قد تأخذ الصورة الكمية او الكيفية ، وفى انها تتحدد فى ضوء ما يجب أن يكون عليه الأداء الاختبارى وليس ما هو عليه بالفعل (كما هو الحال فى المعايير) ٠

ولعل هذا النوع هو ما يقصده بناة ما يســـمى فى الوقت الحاضر الاختبارات المرجعة الى المحك ، « فالمحك » المستخدم فى هذه الاختبارات ليس اساسا « خارجيا » ، وانما هو أساس « داخلى » من بنية الاختبار ذاته ولعل أول مع أوقع حركة الاختبارات النفسية فى هذا الخلط روبرت جليزر R. Glaser ، أول من استخدم هذا المصطلح عام ١٩٦٣ دون أن يتنبه الى استخدام مصطلح « محك » فى القياس النفسى بمعنى ثابت ومحدد ٠ ومما يؤسف له حقا أن سار على منواله عدد كبير من قادة القياس النفسى أنفسهم ، وحينما وجدوا أنفسهم فى حيرة من امرهم حاولوا البحثعن حلول توفيقية لعل أشهرها محاولة انستازى فى طبعة ١٩٨٢ من كتابها الشهير (١١ : ٩٤ ـ ١٠١) ٠

ونعرض فيما يلى استخدام المستويات فى أنواع مختلفة من الاختبارات النفسية :

١ ـ **مستوى نطاق المحتوى** : لعل أقوى ما يدعم وجهة نظـــرنا فى استخدام مصطلح مستوى بدلا من مصطلح محك كأساس تفسيرى للاداء فى « الاختبارات المرجعة الى المحك » ما تتطلبه طبيعة هـــذه الاختبارات ومعظمها من نوع الاختبارات التحصيلية من الرجوع ـ فى الحكم والتفسير على الأداء الاختبارى ٠ الى نطاق المحتوى content domain بدلا من الرجوع الى خصائص عينة تقنين معينة (كما هو الحال فى المعيار) أو الى خصائص أداء مستقل عن الاختبار (كما هو الحال فى المحك) ، ولهذا يشار اليها احيانا بالاختبارات المرجعة الى المحتوى أو النطاق أو الأهداف ، وهذا ما تفضله انستازى (١١ : ٩٤) ، لأن مصطلح « الاختبارات المرجعة

الى المحك » - على الرغم من شيوعه - ليس أفضل ما يشير الى هذا النوع من الاختبارات ·

ويتم الحكم على الأداء الاختباري وتفسيره هنا في ضوء معنى المحتوى حيث يتم التركيز على ما يؤديه المفحوص بالفعل مقارنا بما يجب أن يؤديه ، وليس مقارنا بأداء الآخرين · ويتطلب ذلك أن يتحدد مقدما نطاق المعرفة أو المهارة التي يقيسها الاختبار · ومن المطالب الجوهرية في هذه الحالة أن يكون » نطاق المحتوى « هذا الذي يتم اختياره - أو عينة منه - في بناء الاختبار له أهمية ، ثم يقسم النطاق المختار الى وحدات صغيرة في صورة عبارات أدائية · وهذا ما يسمى في السياق التربوي صياغة الأهداف التعليمية صياغة سلوكية اجرائية (راجع ، ٧)·

والهدف الجوهري للمقارنة في هذه الحالة هو الحكم على مدى الاتقان mastery وهو جوهر الحكم في هذه الحالة ، وهو من نوع الكل أو لا شيء ، ويشير الى مدى وصول المفحوص الى المستوى المحدد مقدما للاتقان ، وقد تستخدم ثلاثة مستويات للحكم على مستوى الاتقان ، المستوى المتوسط أو موضع الشك أو الذي يحتاج لمراجعة ، ومستوى عدم الاتقان أو الفشل ويحتاج هذا بالطبع الى تحديد نقطة القطع cutoff point التي عندها يمكن أن تصدر مثل هذا الحكم · وتوجد طرق مختلفة لتحديد نقاط القطع يدرسها بالتفصيل محمود محمد ابراهيم عبد الله في رسالته للماجستير بكلية التربية جامعة الأزهر (١١) ·

ويحتاج الوصول الى الحكم في ضوء المستوى استخدام اجراءات القرار التتابعي sequential decision وفيها لا يعرض على المفحوص عدد ثابت محدد مقدما من الأسئلة - كما هو الحال في الاختبارات المعتادة ، وانما يستمر المفحوص في اخذ الاختبار حتى يصل الفاحص الى نقطة القطع التي عندها يصل الفاحص الى اتخاذ قرار باتقان المفحوص للمحتوى أو عدم اتقانه له · وعند هذه النقطة يتوقف الاختبار ويوجه المفحوص اما الى المستوى التعليمي التالي او يعود الى تعلم المحتوى نفسه من جديد · وقد اصبحت هذه الاجراءات التتابعية اكثر يسرا مع استخدام الحاسبات

الالكترونية .التى تختصر وقت الاختبار عن ناحية وتوصل الى تقدير للاتقان اكثر دقة ٠

٢ ـ **المعايير الارتقائية الرتبة :** تعود هذه المعايير باصولها الى نتائج بحوث علم نفس النمو والتى أدت الى وصف السلوك الانسانى فى مراحله المتتابعة ومن اشهر الاعمال المبكرة فى هذا المجال جهود جيزل وزملائه فى جامعة ييل بالولايات المتحدة والتى نتجت عنها جـــداول جيزل الارتقائية Gesell Developmental Schedules وتوضح هذه الجداول المســـتوى النمائى التقريبى التى يصل اليه الطفل مقدرا بالشهور فى أربعة مجالات رئيسية هى النشاط الحركى والتوافقى واللغوى والشخصى الاجتماعى ٠ ويتحدد هذا المستوى بمقارنة سلوك الطفل وبالسلوك المعتاد للاطفـــال من سن ٤ أسابيع وحتى ٣٦ شهرا ونشأت عن ذلك مقاييس ذات طبيعة رتبيـة متضمنة ثلاثة افتراضات أساسية هى :

١ ـ يتسم السلوك المبكر فى نموه بالتتابع المنتظم ٠

٢ ـ تتتابع المراحل النمائية الارتقائية لسلوك الانسان بنظام ثابت ٠

٣ ـ تعتمد كل مرحلة لاحقة على اتقان الأنماط الســـلوكية اللازمة للمراحل السابقة ٠

ومنذ مطلع الستينات شهد الميدان اهتماما كبيرا بالنظريات الارتقائية لعالم النفس السويسرى العظيم جان بياجيه ، والتى توصل اليها من خلال برنامج طموح للبحث فى النشاط المعرفى عند الاطفال والمراهقين بداه منذ منتصف العشرينات ٠ وقد وظفت المهام التى استخدمها فى معظم الدراسات اللاحقة فى علم نفس النمو ، كما بنيت اختبارات عقلية مقننة على أساسها وتعد هذه الاختبارات من النوع الرتبى أيضا وتصدق عليها الافتراضـــات السابقة ٠

وللتمييز بين هذا النوع من المعايير ومعيار العمر العقلى نشير الى إنه فى المعايير الرتبية تصمم المهام (التى قد تتحول الى أسئلة أو مفردات فى

الاختبار) بحيث تكشف عن الجوانب الهامة والسائدة لكل مرحلة ارتقائية، ثم تجمع بعد ذلك البيانات الامبريقية حول الاعمار التى يتم التوصل فيها الى كل مرحلة من هذه المراحل • أما فى معيار العمر العقلى فالعكس هو الصحيح ، ففيه تختار المهام (الأسئلة) على اساس تمييزها بين الأعمار المتتابعة • ويوجد وجه آخر للاختلاف بين نوعى المعايير هو أن معيار العمر العقلى كمى فى جوهره ويتحدد فى ضوء درجة يحصل عليها المفحوص فى الاختبار • أما فى المعيار الارتقائى الرتبى فان الأهم هو الوصف الكيفى لسلوك الطفل ولذلك فهو يعتمد على المستوى وليس المعيار •

مراجع الباب الثانى

١ ـ أحمد الرفاعى محمد غنيم : تطبيقات على ثبات الاختبارات ، القاهرة ، مكتبة نهضة الشرق ، ١٩٨٥ .

٢ ـ رمزية الغريب : التقويم والقياس النفسى والتربوى ، القاهرة : مكتبة الأنجلو المصرية ١٩٧٠ .

٣ ـ سيد أحمد عثمان : لمحة الى ما وراء المنهج . مقدمة الطبعة الثانية لكتاب (مناهج البحث فى التربية وعلم النفس) تاليف د.ب. فان دالين (ترجمة محمد نبيل نوفل وآخرين) مكتبة الأنجلو المصرية ، ١٩٧٩ .

٤ ـ ـــــــــــــ : الموضوعية والذاتية . الكتاب السنوى فى التربية وعلم النفس (تحرير سعيد اسماعيل على) ، المجلد الخامس ، ١٩٧٨ .

٥ ـ صلاح قنصوة : الموضوعية فى العلوم الانسانية ، دار الثقافة للطباعة والنشر ، ١٩٨٠ .

٦ ـ صلاح مخيمر : عن الذاتية والموضوعية فى علم النفس ، مكتبة سعيد رأفت ، (ب.ت) .

٧ ـ فؤاد أبو حطب ، آمال صادق علم النفس التربوى (ط ٣) ، مكتبة الأنجلو المصرية ، ١٩٨٤ .

٨ ـ فؤاد أبو حطب قضية سيرل بيرت ، محاضرة عامة ألقيت بالجمعية المصرية لدراسات النفسية ٢ سبتمبر ١٩٧٨ .

٩ ـ فؤاد أبو حطب ، بديوى علام : الاحصاء النفسى والتربوى . مكتبة الأنجلو المصرية (تحت الطبع) .

١٠ ـ محمد خليفة بركات : مشروع ميدنى لميثاق أخلاقيات المهنة للمشتغلين

بعلم النفس - المؤتمر السنوى الثانى لعلم النفس ، الجمعية المصرية للدراسات النفسية ، ابريل ١٩٨٦ .

١١ - محمود محمد ابراهيم عبد الله : دراسة مقارنة لبعض طرق حساب معامل الاعتمادية لنقاط القطع فى الاختبارات مرجعية المحك . رسالة ماجستير - كلية التربية جامعة الازهر ١٩٨٦ .

11. Anastasi, A. **Psychological Testing.** (3rd. ed.) New York : Macmillan, 1968, 1976, 1982.
12. A.P.A. **Ethical Standards** for psychoogists. **Washington,** D.C., 1953.
13. A.P.A. **Technical Recommendations** for psychological **Tests and Diagnostic** Techniques, Washington D.C. 1954.
14. A.P.A. Standards for educational and psychological tests and **manuals. Washington, D.C. 1966, 1974.**
15. A.P.A. Ethical Standards for Psychologists Amer. Psychologist, 1968, 23, 357-361.
16. ———— Casebook on ethical statements of psychogists. Washington, D.C., 1974.
17. ———— Ethical principles in the conduct of research with human participants, Washington, D.C., 1973.
18. Benton, A. L. Influence of incentives upon intelligence test scores of school children, J. **Genet. Psychol.,** 1936, 49, 494, 496.
19. Borden, E.S. **Psychological Counseling,** New York, Appleton, 1965.
20. Cronback, J.L. **Essentials of Psychological Testing** (3rd ed.) New York : Harper, 1960, 1970, 1984.
21. Cronbach, J.L., Meehl, P.E. **Construct validity in psychological tests. Psychol. Bull., 1955, 52, 281-302.**
22. Flanagan, J.C. The development of an index of examine motivation **Educ. Psychol. Measurement, 1955, 15, 144-161.**
23. **Ghiselli, G.E., Campbell, J.P. and Zadeck, S. Measurement theory for the behavioural sciences, San Francisco : W.H. Freeman, 1981.**
24. **Guilford, J.P. Psychometric methods. New York : McGraw-Hill, 1954.**
25. **Hurlock, E.B. An evaluation of certain incentives used in school Work. J. Educ. Psychol., 1925, 16, 154-159.**

26. Lindquist, E.F. (ed.) Educational measurement Washington, D.C., Amer. Council on Educ., 1951.
27. Nunnally, J.C. Psychometric theory. New York : McGraw-Hill, 1978.
28. Ronan, W.W. & Prien, E.P. Perspectives on the measurement of human performance. New York : Appleton-Century-Crofts, 1971.
29. Sarason, S.B. et al., A test anxiety Scale for children. Child Develop., 1959, 29, 105-115.
30. Stanley, J.C. & Hopkins, K.D. Educational and Psychological measurement evaluation. N.J. : Prentice-Hall, 1972.
31. Thorndike, R.L. (ed.) Educational Measurement (2nd ed.), Washington : D.C. Amer. Council on Educ., 1971.
32. Thorndike, R.L. & Hagen, E. Measurement and evaluation in psychology and education (4th ed.) New York : Wiley, 1977.

البـــاب الثالث

اختبارات الذكاء العـــام

تقـــديم للباب الرابع

توجد فئة من الاختبارات النفسية تسمى اختبارات الذكاء العام · وتتميز هذه الاختبارات بأنها تستخدم فى مواقف متنوعة ويتحدد صدقها فى ضوء مد كاتأكثر شمولا · كما تتميز هذه الاختبارات بأنها تعطى درجة واحدة مثل نسبة الذكاء لتدل على المستوى العقلى العام للمفحوص · ويرى ترمان وميرل (٥٦) أن الوصول الى هذا التقـدير الشمولى للمســتوى العقلى يتطلب خبرة من الباحث فى «تصويب الرماح الى الأهداف الحاسمة »· ولتحقيق ذلك عليه أن يعرض على المفحوص عددا كبير ومتنوعا من الأعمال على أساس أنها تمثل عينة ملائمة من أساليب الأداء العقلى · ومع ذلك فان البحوث الحديثة تؤكد أن هذه الاختبارات تتشبع بعمليات عقلية معينة مثل القدرة اللغوية ، وتتجاهل تماما بعض العمليات العقلية الأخرى ، كالقدرة الموسيقية ·

وكثيرا ما تسمى اختبارات الذكاء العام باختبارات الاستعداد المدرسى Scholasic aptitude لان صدقها يتحدد فى العادة فى ضـــوء محكات التحصيل الأكاديمى · كما ان هذه الاختبارات تستخدم فى أغلب الأحوال فى أغراض التصفية المبدئية العامة (سواء من الناحية التربوية أو المهنية او العسكرية) ، والمفروض ان يتبعها تطبيق اختبارات الاستعدادات الأكثر تخصصا · ويشيع هذا على وجه الخصوص فى اختبار المراهقين والراشدين الأسوياء عند الارشاد أو الانتقاء أو التصنيف أو التوجيه أو غير ذلك من الأغراض التى اشرنا اليها فى الفصل الأول من هذا الكتاب · كما يوجــد استخدام آخر لاختبارات الذكاء ، فى الأعراض الكلينيكية خاصة ، وهو تحديد الضعف العقلى وتشخيصه · وفى مثل هذه الأحوال فان استخدام الاختبارات الفردية أكثر شيوعا من غيره ، الا اننا كثيرا ما نســــتخدم الاختبارات الجماعية للذكاء فى أغراض البحث والممارسة السيكووجيتين ·

وسوف نخصص فى هذا الباب لاختبارات الذكاء العام فيتناول الفصل السابع اختبارات الذكاء الفردية أما الفصل الثامن فسوف يتبادل الاختبارات الجماعية للذكاء ·

الفصل السابع

اختبارات الذكاء الفردية

يتناول هذا الفصل بالعرض اختبارات الذكاء الفردية ، وعلى وجه الخصوص اشهر اختبارين فرديين للذكاء وهما : مقياس ستانفرد ـ بينيه ، ومقياس وكسلر ، ويعد مقياس بينيه على وجه الخصوص الوالد الشرعي لحركة قياس العقل الانسانى ـ وخاصة عند الأطفال ، ولذلك سيحظى فى هذا الفصل باهتمام خاص ، ثم نعرض لمقاييس وكسلر ، ولبعض الاختبارات والمقاييس الأخرى الأقل شهرة والتى لا تقل أهمية ٠

مقياس ستانفرد ـ بينيه

الفرد بينيه :

الفرد بينيه (١٨٥٧ ـ ١٩١١) هو العالم الفذ الأصيل ، بدأ تدريبه فى ميدان الطب ثم سرعان ما أصبح أشهر علمـــاء النفس فى فرنسا فى عصره ؛ وله فضله العظيم على علم النفس فى بلده ٠ فقد أنشأ مع زميله بيونى Beauns أول معمل لعلم النفس عام ١٨٨٩ فى جامعة السوربون ، وأول مجلة متخصصة باللغة الفرنسية « حوليـــة علم النفس L'Année Psychologique عام ١٨٩٥ ٠

ولقد بدأ بينيه عالما تجريبيا ، فقبل أن يعد اختباراته العقلية المشهورة اهتم بدراسة العمليات العقلية المعرفية ، واطلق على منهجه اسم «سيكولوجية الفرد » ٠ وفى رأيه أن أهم ما نجده من فوارق بين الناس انما يكون هو فى الأنشطة العقلية العليا كالاستدلال والحكم ، وهى جميعا من العمليات التى يصعب قياسها بالوحدات الفيزيائية (كالاحساس) ، كما لا تتوافر فيهـا الوحدات الأولية للنشاط الذهنى ، والتى كان يهتم بها أغلب علماء النفس من أنصار بنائيه فوندت (ومنهم ، تتشنر) (١٤) ٠

وقد تأكدت صحة وجهة نظر بينيه من نتائج بعض الدراســات التى

اجريت على « صدق اختبارات جيمس ماكس كاتل (الذى اقتفى فيها اثر السير فرنسيس جالتون فى انجلترا) لمقاييس عقلية ، وبخاصة بحوث كلارك وزلر وشتيلا شارب (٢٤) ، ولذلك بدأ بينيه يطبق فى بحوثه عن الاختبارات العقلية ما توصل اليه من نتائج البحث التجريبى فى عملية التفكير ، فانتقد اختبارات كاتل وجالتون لأنها تقتصر على النواحى الحسية البسيطة ، وفضل الاختبارات الأكثر تعقدا وتركيبا ، واقترح عددا من الاختبـــارات لقياس « الملكات » العقية (٥٠) ٠

وقضى بينيه الفترة بين عامى ١٨٩٦ و ١٩٠٥ فى اجراء سلسلة طويلة من البحوث على الأطفال الأسوياء وغيرهم مطبقـــا أنواعا عـــديدة من الاختبارات لقياس مختلف « الملكات » العقلية ، بالاضافة الى عـــدد من التجارب أجراها حول المسائل الخلافية فى سيكولوجية التفكير فى ذلك الوقت مثل التجريد ، والتفكير بدون صور وزمن الرجع والتمييز الحسى ٠ وكان الهدف الذى يسعى اليه من ذلك هو تحديد العلاقة بين هذه الاختبارات والعمر الزمنى والتحصيل المدرسى وتقديرات المعلمين للقدرة العقلية ، وكانت بحوثه جيدة التصميم وأدت به الى التحول من الاهتمام بالاختبارات التى تقيس « الملكات » التقليدية الى تلك التى تقيس المستوى العقلى بالفعل ٠

وهكذا كانت نظرة بينيه الى الذكاء أكثر شمولا مما ظهر فى مقياسه ذاته بعد ذلك ؛ حين لجأ الى استخدام درجة كلية واحدة هى العمر العقلى، فى قياس الذكاء ٠ ويمكن للقارئ المهتم باســهامات بينيه فى النظـــرية الأساسية للذكاء أن يرجع الى (٢٤) ٠

نشأة مقياس بينيه :

فى عام ١٩٠٤ طلبت وزارة المعارف الفرنسية من الفرد بينيه والطبيب الفرنسى ثيودور سيمون دراسة الطرق التى يمكن أن تســتخدم فى تصنيف الأطفال الذين يتميزون ببطء التعلم فى المدارس الفرنسية ، فعكفا على وضع مقياسهما المشهور للذكاء بغرض التمييز بين الأطفال الأســوياء وضعاف العقـــل ٠

ولم يكن الهدف دراسة وتحليل استعدادات المتخلفين عقليا ، وانما كان الهدف عمليا بحتا وهو تقويم وقياس ذكاء هؤلاء الاطفال بصفة عامة وتحديد مستواهم العقلى ومقارنتهم بالأطفال الأسوياء من نفس العمر والمستوى · أى الهدف كما يقول بينيه وسيمون (٥٠) هو بناء مقياس « مترى » Metrical للذكاء يتكون من عدد من « الاختبارات » المتدرجة في الصعوبة ، مرتبة حسب الأعمار حتي يستطيع الفاحص أن يحدد مدى تقدم الطفل عقليا أو تخلفه · وقد سعى بينيه الى أن تكون هذه « الاختبـــارات » متنوعة ، مع التركيز على فئات معينة من العمليات العقلية اهتم بها دون غيرها عند اعداد المقياس وهى : الحكم والفهم العام والمبادأة والقدرة على التكيف ·

ويتكون مقياس بينيه الأصلي الذى نشر عام١٩٠٥باسم مقياس بينيه ـ سيمون للذكاء Binet-Simon Scale من ٣٠ سؤالا أطلق على كل منهـا اسم « اختبار » وشاع هذا الاسم فى جميع الطبعات التالية لهذا المقياس فى جميع اللغات (*) ، وفيما يلى قائمة بهذه الاسئلة ، نذكرها كاملة لنبين الى أى حد اختلفت عن اهتمامات علماء القياس العقلى قبل بينيه ، وكيف أنها هيأت الموقف لظهور عدد كبير من الاختبارات العقلية تنحدر فى معظمهـا « من صلب » مقياس بينيه على حد تعبير أنستازى (٤٩) بالاضـــافة الى شيوعها فى معظم تعديلات مقياس بينيه ذاته ·

(١) التآزر البصرى (تتبع شمعة مضيئة بالرأس والعينين) ·
(٢) الفهم عن طريق ولمس (الامساك بمكعب بعد لمسه) ·
(٣) الفهم عن طريق البصر أو الرؤية ·
(٤) التعرف على الطعام (التمييز بين قطعةالخشبوقطعة الشيكولاته)·
(٥) البحث عن الطعام (الاستجابة لقطعة من الشيكولاته ملفوفة فى الورق) ·
(٦) تتبع التعليمات البسيطة أو الايماءات المتكررة ·

(*) آثرنا أن نشير الى « الاختبارات » التى يتألف منها مقياس ستانفرد ـ بينيه بمصطلح « الاسئلة » فى متن هذا الكتاب ، حتى لا تختلط بمعنى كلمة « الاختبار » كما تشيع فى التراث السيكولوجى الحديث ·

(٧) الاشارة الى الأشياء وخاصة أعضاء الجسم (الرأس ، الأنف ، الخ) ٠

(٨) التعرف على الأشياء من صورها ٠

(١٠) التمييز الحسى بين طول خطين ٠

(١١) تكرار ثلاثة أرقام ٠

(١٢) التمييز بين وزن شيئين ٠

(١٣) القابلية للايحاء المقارنة بين خطين متساويين ـ أو البحث عن أشياء لا وجود لها فى صورة) ٠

(١٤) تعريف الكلمات البسيطة (التى تدل على أشياء مألوفة) ٠

(١٥) تكرار جملة تتكون من ١٥ كلمة ٠

(١٦)بيان أوجه الاختلاف بين شيئين ٠

(١٧) الذاكرة البصرية ٠

(١٨) رسم الأشكال من الذاكرة ٠

(١٩) مدى ذاكرة الارقام ٠

(٢٠) بيان أوجه التشابه بين الاشياء ٠

(٢١) التمييز بين الخطوط بسرعة ٠

(٢٢) ترتيب ٥ أوزان ٠

(٢٣) تحديد الوزن الناقص (من بين الاوزان الخمسة فى الاختبار السابق) ٠

(٢٤) اعطاء كلمات ذات قافية واحدة ٠

(٢٥) تكملة الجمل ٠

(٢٦) بناء جملة تحتوى على ٣ كلمات معينة ٠

(٢٧) الاجابة على الأسئلة (مثل : ماذا تفعل وانت نعسان ؟) ٠

(٢٨) تحديد الوقت بعد أن تحل عقارب الساعة محل بعضها ٠

(٢٩) ثنى وتقطيع قطعة من الورق ٠

(٣٠) التمييز بين الكلمات المجردة (مثل ، حزن وسأم) ٠

واذا فحصنا هذه القائمة فحصا دقيقا نجد أن الاسئلة الثلاثة الأولى هى فى الواقع أسئلة للنمو الحركى ، أما الاسئلة الأخرى فيمكن ان تعد من النوع العقلى ، ومن بينها ١٨ سؤالا (أى حوالى الثلثين) من النوع الذى يقيس ما يسميه جيلفورد التفكير المعرفة و ٥ أسئلة تقيس ما يسميه الذاكرة

بانماطها المختلفة (الاسئلة ١١ ، ١٥ ، ١٧ ، ١٨ ، ١٩) ولا يوجد الا ثلاثة اسئلة فقط تقيس ما أطلق عليه فيما بعد قدرات التفكير الابتكارى (او الانتاج التباعدى عند جيلفورد) ٠

وكانت الطريقة التى استخدمها بينيه فى تقنين مقياسه غير ملائمة ، فقد طبقه على ٥٠ طفلا (١٠ من كل فئة من فئات الاعمار ٣ ، ٥ ، ٧ ، ٩ ، ١١) اختارهم معلموهم على أساس أنهم متوسطو القدرة العقلية ، بالاضافة الى عدد من الاطفال لم يذكره بينيه كانوا من ضعاف العقول ٠ وكانت طريقة التصحيح تعتمد على أساس تجريبى مبدئى يسمح بتصنيف الاطفال تبعـا لأعلى مستوى يصل اليه الطفل : فالمعتوهون لا يمكنهم تجـــاوز السؤال السادس ، ولا يتجاوز البلهاء السؤال الخامس عشر ، أما حدود المأفونين فلم تكن معروفة ، رغم أنهم يعجزون فى العادة عن الاجابة على الأسئلة المجردة التى يستطيع معظم الاطفال العاديين من سن ١١ الاجابة عليها ٠

وبالطبع ارتبط عمر المفحوصين بنجاحهم فى الاختبار ، ولم تكن لدى بينيه حتى ذلك الوقت (عام ١٩٠٥) طريقة تسمح بالمقارنة بين الأعمار ٠ ولو انه كان يلتمس طريقه فيما يبدو نحو مفهوم العمر العقلى ٠ يقول بينيه وسيمون عام ١٩٠٥ (٥٠ : ٢٢١) ٠

« ومن المحتمل أن ننجح فى يوم ما فى ايجاد علامات على التخلف السيكولوجى مستقلة تماما عن العمر ٠ وقد يكون من المفيد جدا التعرف على هذه العلامات ، الا أننا فى الوقت الحاضر نجد ان ما يجذب انتباهنا أثر من غيره هو التشابه بين الأسوياء من صغار السن وغير الأسوياء من كبار السن ٠ وهذا التشابه من الكثرة والغرابة بحيث اننا لو حصـلنا على استجابات طفل لا نعرف عمره لا نستطيع الحكم عليه بالســـواء أو عدم السواء ٠

وكانت لطبعة عام ١٩٠٥ من مقياس بينيه عدة سمات تعد نقطة تحول فى تاريخ القياس العقلى ففيه نجد لأول مرة ان الأسئلة التى تقيس مكونات عديدة ترتبط فيما بينها لتحدد مستوى من الأداء بدلا من القياس المنفصل لأبعاد التحليل السيكولوجى التقليدى الذى كان شائعا فى عصر بينيه مثل

الملكات أو عتبات الاحساس أو غيرها وبالاضافة الى ذلك فان وضع سؤال فى المقياس أو حذفه أو تعديله كان من المسائل التجريبية ، ويدل على خبرة بينيه الطويلة بالتجريب العملى . فقد كان يهتم بمدى العلاقة بين السؤال وأحكام المحك (أى بنوع من الصدق) . ثم ان هذه الاختبـــارات لم تكن محددات معملية دقيقة وانما كانت أدوات عملية بسيطة وهامة لدراسة الأطفال فى مواقف واقعية (مثل حجرة الدراسة) .

وقد أدى ظهور مقياس ١٩٠٥ الى جذب انتباه عدد كبير من المهتمين بمشكلات الضعف العقلى والتصنيف التربوى ، فقـــام العالم الأمريكى جودارد Goddaard بترجمته الى الانجليزية ، كما جـــربه ديكرولى Decroly وديجان Degand فى بلجيكا . وكان الكثيرون ـ ومنهم بيرت فى انجلترا وترمان فى أمريكا ـ على اتصال مباشر ببينيه كما اقترحوا فى نفس الوقت اختباراتهم الخاصة .

تعديل عام ١٩٠٨ :

فى عام ١٩٠٨ ظهر مقياس جديد لبينيه وسيمون يعتمد اعتمادا كبيرا على المقياس السابق ومع تعديلات جوهرية ، فقد أصبح اسم المقياس الجديد « نمو الذكاء عند الأطفال» The development of intelligence in children وتحول تركيز هذا المقياس من غير الأسوياء الى الأسوياء ، واســـتبعدت اختبارات المعتوهين . واحتفظ المقياس بخاصية استخدام الأسئلة القصيرة، العديدة المتنوعة من حيث المحتوى ، والمرتبة ترتيبا تجريبيا حسب الصعوبة . ولأول مرة تصنف الأسئلة حسب مستويات الأعمار ابتداء من سن ٣ سنوات حتى سن ١٣ سنة تبعا للعمر الذى يستطيع الأطفال الأسوياء الوصول الى أسئلته بنجاح . وكان تحديد الأسئلة للمستويات العمرية المختلفة تحديدا جزافيا ، ليس بسبب قلة عدد أفراد عينة التقنين فحسب وانما أيضا بسبب الصعوبات العملية فى اعداد الأسئلة التى تقع فى مستويات الصـــعوبة المطلوبة ، وبوجه عام اعتبر المستوى العمرى للاختبار هو ما يستطيع ما بين ٥٥٪ ـ ٧٥٪ من أطفال عينة التقنين بالاجابة عليه اجابة صحيحة . لأن هذا المدى يتضمن نسبة الـ ٥٠٪ التى تدل دلالة عامة على الأداء المتوسط وكانت عينة التقنين ٢٠٣ أطفال تتراوح أعمارهم بين ٣ سنوات و ١٣ سنة .

وقد أدى ترتيب الأسئلة النمائى حسب مستويات العمر بالفاحصين الى التعبير عن المستوى الارتقائى للطفل بالعمر الذى يحصل فيه الطفل المتوسط على درجة مكافئة . ورغم أنه لم تتحدد مستويات النجاح والفشل على اساس أسئلة المقياس ، فان العمر العقلى للطفل كان يتحدد على اساس يساوى العمر الذى يستطيع أن يجتازه ويجتاز الأعمار السابقة عليه (العمر القاعدى أو الاساسى) ثم يضاف اليه عام اضافى من العمر العقلى لكل ٥ أسئلة من المستويات الأكثر صعوبة .

هذه هى الخصائص الأساسية لمعيار العمر العقلى كما اقترحه تعديل ١٩٠٨ ، وهو المعيار الذى اقتبسته مقاييس الذكاء اللاحقة . ورغم أن هذا المعيار يتضمن مشكلات احصائية أشرنا اليها فى الفصل السادس ، فان له فائدته فى أنه يوفر للباحث طريقة سهلة الفهم فى وصف الفروق بين الأطفال فى المستوى العام للقدرة العقلية . ومع ذلك بقيت مشكلتان هامتان لم يحلهما مقياس ١٩٠٨ : أولاهما أن عدد الأسئلة فى كل مستوى عمرى يختلف ، ويمتد بين ٣ ، ٨ أسئلة ، وثانيهما أن درجة العمر العقلى لا تفهم كمعيار للتفوق أو التخلف الا اذا ارتبطت بالعمر الزمنى للطفل .

تعديل عام ١٩١١ :

فى عام ١٩١١ قام بينيه - قبيل وفاته المبكرة - بتعديل مقياسه ونشره باسمه منفردا ، فأعاد ترتيب كثير من الأسئلة ، ووحد عددها فى كل مستوى عمرى فجعله خمسة أسئلة ثم أضاف عددا آخر من الاسئلة (٥ لمستوى ١٥ سنة ، و ٥ لمستوى الراشد) وهكذا أمكن - بعد تساوى عدد الاختبارات فى كل مستوى - اعطاء كسور من العمر العقلى مقدارها سنة لكل اختبار يستطيع المفحوص الاجابة عليه بعد عمره القاعدى أو الأساسى . ومع ذلك فان مفهوم العمر العقلى ظل على حاله ، أى لا يمكن استخدامه كمؤشر على المستوى العقلى مستقلا عن العمر الزمنى ، وقد أدى ذلك بالعالم الالمانى شترن Stern عام ١٩١٢ الى اقتراح معيار النسبة العقلية أى $\frac{\text{العمر العقلى}}{\text{العمر الزمنى}}$.

وهو المعيار الذى استخدمه ـ بعد التعديل ـ ترمان كما سنوضح بعد قليل ٠
ويوضح الجدول رقم (٨) الأسئلة التى يتكون مقياس عام ١٩١١ ، ومن هذا الجدول يتضح أن المقياس فى صورته الأخيرة على يد بينيه يتكون من ٥٤ سؤالا ، وفيه حذفت أسئلة المستويات الخاصة بالأعمار ١١ ، ١٢ ، ٢٤ سنة كما أن أسئلة سن ٤ كانت ٤ أسئلة فقط (٥٣) ٠

جدول رقم (٨) يبين الاختبارات التي يتكون منها مقياس بينية (تعديل عام ١٩١١)

العمر	أسئلة تعديل عام ١٩١١				
٣	اسم العائلة	الإشارة	عددان	٦ مقاطع	صورة (سهل)
٤	الجنس	التسمية	٣ أعداد	خطان	
٥	٤ قطع نقود	المربع	وزنان	١٠ مقاطع	بطاقة مقسمة
٦	١٣ قطعة نقود	المعين	مقارنة الوجوه	تعريف الكلمات (سهل)	الصباح والمساء
٧	٩ قطع نقود، ٣ أنصاف لهذه القطع	٤ ألوان	اليمين والشمال	الترتيب الثلاثي	صورة (صعب)
٨	السمات الناقصة	التاريخ	٥ أعداد	الفروق (المحسوسة)	العد من ٢٠ إلى ١
٩	٩ قطع نقود	التغيير	الشهور	تعريف الكلمات (محسوس)	الأسئلة السهلة
١٠	بناء الجمل (سهل)	الرسم من الذاكرة	٥ أوزان	السخافات	الأسئلة الصعبة
١٢	بناء الجمل (صعب)	الإيحاء	٦٠ كلمة في ٣ دقائق	تعريف الكلمات (مجرد)	الجمل المختلطة
١٥	المشكلات والمسائل	٣ إيقاعات	٧ أعداد	٢٦ مقطعاً	الصورة
الرشد	الورق المطوى	إعادة التعبير	الفرق بين حرفين	الفروق المجردة	المثلث المعكوس

مقياس ستانفرد ـ بينيه لترمان :

لم تكن محاولة جو دارد ترجمة ونقل مقياس بينيه الى اللغة الانجليزية هى المحاولة الوحيدة ، ففى عام ١٩١٠ نشر هوى Huey وويبل Whipple ترجمة له ، كما نشر فالين Wallin وكولمان Kuhlmann ترجمة اخرى فى عام ١٩١١ . وفى عام ١٩١٣ بدا بيرت Burt نقل المقياس الى الانجليزية وتقنينه على عينة كبيرة من الاطفال الانجليز عندما عين اخصائيا نفسيا لمقاطعة لندن ، وظل يتابع مشروعه هذا طوال تاريخه العلمى كأستاذ لعلم النفس بجامعة لندن . بالاضافة الى نقل المقياس فى اقطار عديدة بلغاتها مثل ايطاليا وألمانيا وروسيا والسويد وتركيا واليابان . وفى عام ١٩٢٠ ترجم أحمد فكرى (٥) مقياس بينيه الأصلى الى اللغة العربية . وقدم له تعليمات مقتضبة ، وكان يعتقد خطأ أن المدرس ـ دون تدريب سابق ـ يمكنه تطبيق المقياس والحكم على ذكاء الطفل . ومع ذلك فان اكثر هذه المحاولات شهرة فى تاريخ القياس النفسى محاولة لويس ترمان L.M. Terman (١٨٧٧ ـ ١٩٥٦) أستاذ علم النفس بجامعة ستانفرد . والتى عرف بها المقياس باسم مقياس ستانفرد ـ بينيه للذكاء ، وقد شاعت هذه التسمية اكثر من الاسم الأصلى للمقياس .

وقد ظهرت الطبعة الأولى لمقياس ستانفرد ـ بينيه عام ١٩١٦ . وادخل ترمان على المقياس الأصلى عددا من التعديلات والاضافات الجوهرية بحيث يمكن القول أنه اختبار جديد . فقد كان حوالى ثلث الأسئلة جديدا ، وعدل ترمان عددا من الأسئلة القديمة واعيد توزيعها على المستويات العمرية المختلفة أو حذفت . واعيد تقنين المقياس كله على عينة من الأمريكيين بلغ عددها ١٠٠٠ طفل و ٤٠٠ راشد . واعدت تعليمات مفصلة لتطبيق كل سؤال وتصحيحه ، واستخدام معيار نسبة الذكاء (تعديلا لمعيار النسبة العقلية لشترن وكولمان) فى صورتها التى شاعت فى القياس النفسى وهى :

$$\text{نسبة الذكاء} = \frac{\text{العمر العقلى}}{\text{العمر الزمنى}} \times ١٠٠$$

وقد كتبت مجلة المقتطف المصرية مقالا عن هذا المقياس عام ١٩٢٧(٤٦) كما كتب عنه أمير بقطر عام ١٩٢٨ (١٠) . ونقله الى اللغة العربية لأول مرة طبيب غير متخصص فى علم النفس هو حسن عمر (١٥) . ثم ترجمه على نحو أدق اسماعيل القبانى (٦) عام ١٩٣٨ .

وفى عام ١٩٣٧ ظهرت فى الولايات المتحدة الطبعة الثانية المعـــدلة للمقياس ، وتتكون من صورتين متكافئتين ، الصورة ل ، والصورة م . وفى هذه الطبعة زاد عدد أسئلة المقياس وأعيد تقنينه تماما على عينة جديدة جيدة السحب من المجتمع الأمريكى بلغ عـــددها اكثر من ٣٠٠٠ مفحوص تتراوح أعمارهم بين سن ٥ر١ سنة الى سن ١٨ . واشتملت كل عينة عمرية على عدد متساو من الذكور والاناث ، وتم اختبار معظم المفحوصين من سن ٦ سنوات وما بعدها من تلاميذ المدارس ، أما عينات أطفال ما قبل المدرسة (عينات الأعمار ٥ر١ ، ٢ ، ٥ر٢ ، ٣ ، ٥ر٣ ، ٤ ، ٥ر٤ ، ٥ ، ٥ر٥ سنة) فتم الحصول عليها بطرق مختلفة ، أيسرها تطبيق المقياس على اخوة أطفال وتلاميذ المدارس ، وقد روعى فى اشتقاق العينات التوزيع الجغـــرافى والمستويات الاقتصادية والاجتماعية .

وقد ظهرت الطبعة التجريبية للصورة ل باللغة العربية ١٩٥٦ وقام باعدادها محمد عبد السلام أحمد ولويس كامل مليكه (٤٠) .

وفى عام ١٩٦٠ ظهرت طبعة ثالثة لمقياس ستانفرد - بينيه فى صورة واحدة فقط تجمع افضل الأسئلة من صورتى طبعة عام ١٩٣٧ . وبدون تقديم أى محتوى جديد أو صياغة اسئلة جديدة كان من الممكن حذف الأسئلة التى لا تساير العصر . واعيد توزيع الأسئلة التى تغيرت مستويات صعوبتها فى الفترة بين التعديلين نتيجة للتغيرات الثقافية ، وفى اعداد طبعة عـــام ١٩٦٠ واجه ترمان مشكلة شائعة فى القياس النفسى وتتلخص (٤٩) فى أنه من المستحب عامة ادخال تعديلات دورية على المقاييس للافادة من التقدم العلمى فى ميدان بناء الاختبارات ، ومن الخبرة السابقة باستخدام الاختبار، بالاضافة الى تحديث محتوى الاختبار على نحو يجعله مسايرا لروح العصر . وتزداد الحاجة الى الاعتبار الأخير بالنسبة لأسئلة المعلومات والمواد المصورة والتى تتغير تبعا لتغير الظروف الثقافية . فاستخدام اســـئلة قديمة يؤثر

فى الصدق الظاهرى للاختبار دون شك ، كما قد يغير فى مستوى صعوبة الأسئلة . ومن ناحية أخرى فان التعديل قد يجعل معظم البيانات المتوافرة من البحوث السابقة على الاختبار لا ترتبط بالصورة الجديدة ، فالاختبارات التى استخدمت استخدامات واسعة لمدة سنوات ، كمقياس ستانفرد - بينيه ، تتوافر عنها ثروة من المعلومات يجب أن توزن فى مقابل الحاجة الى التعديل. ولهذه الاسباب فضل تلاميذ ترمان الجمع بين الصورتين فى صورة واحدة بعد انتقاء أفضل ما فيهما . وفقدان صورة مكافئة ليس ثمنا فادحا لتحقيق هذه الغاية ، وخاصة أنه أصبحت توجد منذ عام ١٩٦٠ اختبارات فردية أخرى جيدة الصنع لم تكن متوافرة عام ١٩٣٧ .

وفى عام ١٩٧٢ اجرى تقنين جديد للصورة م - ل باشراف روبرت ثورنديك (٨٠) دون ادخال تعديلات تذكر على محتوى الاختبار ، أما التعديل الجوهرى فقد طرأ على المعايير التى اشتقت من اداء عينة جديدة بلغت ٢١٠٠ مفحوص تم اختبارهم خلال العام الدراسى ١٩٧١ - ١٩٧٢ . وحتى تكون لهذه العينة خصائص التمثيل رغم الصعوبة العمليـة فى تطبيق الاختبارات الفردية لجأ ناشرو الاختبار الى عينة أكبر بلغت مائتى الف تلميذ وطالب من الصفوف من الثالث حتى الثانى عشر (فى النظـام التعليمى الأمريكى) استخدمت فى تقنين أحد الاختبارات الجماعية المشهورة (اختبار القدرات المعرفية الذى سنشير اليه فى الفصل الثامن) ومثلت خصائص المجتمع الأمريكى من حيث الحجم والمناطق الجغرافية والمستويات الاقتصادية الاجتماعية والأصول العنصرية .

وفى اختيار عينة تقنين مقياس ستانفرد - بينيه من بين هذه العينة الأكبر اعتمد على درجات المفحوصين فى البطارية اللفظية فى اختبار القدرات المعرفية المشار اليه بحيث يصبح توزيع درجات العينة الجديدة مطابقا للتوزيع العام على مستوى العينة الكلية . ولم يستبعد الا اولئك الذين لا يتحدثون اللغة الانجليزية . وحتى تشمل العينة الأعمار من ٢ - ٨ سنوات اختـار الباحثون اخوة المفحوصين الذين تضمنتهم العينة وبالمثل اختبرت بعض الحالات الاضافية فى الأعمار العليا . وهكذا أصبحت عينة تقنين المقياس مؤلفة من مائة مفحوص تقريبا لكل فئة عمرية مداها نصف سنة للاعمار من

٢ - ½ سنة ، ومائة مفحوص تقريبا ايضا لكل فئة عمرية مداها سـنة للاعمار من ٦ - ١٨ سنة ٠

وفى عام ١٩٨٥ صدرت الطبعة الرابعة من المقياس من اعداد روبرت ثورنديك واليزابيث هاجن وجيروم ساتلر (٨١) ٠ وقد حافظت هذه الطبعة الجديدة على المعالم التاريخية للمقياس التى استمرت فى الطبعات السابقة ، فأبقيت الأسئلة الصالحة من طبعتى ١٩٣٧ ، ١٩٦٠ ، واضيفت اسئلة أخرى جعلته مقياسا حديثا من حيث المحتوى والخصائص الاحصائية واجراءات التطبيق ٠ واستخدمت فى المقيـاس الجديد أحدث النمـاذج السيكومترية وخاصة نظرية السـمات الكامنة (أو نظرية الاسـتجابة على المفردة item response theory ، وحلل تحيز الاسئلة سواء من حيث الصياغة أو من حيث الخصائص الاحصائية ٠

ولعل اهم التعديلات التى ادخلت على المقياس فى هذه الطبعة استخدام مفاهيم القياس التكيفى للقدرة adaptive ability وذلك بتطبيق نظام الدرجات المتعددة وليس الدرجة الواحدة كما كان الحال فى هذا المقياس طوال تاريخه وأمكن بهذه الطريقة اعطاء المفحوص أربعة انواع من الدرجات خلال زمن الاختبار هى : الاستدلال اللفظى ، والاستدلال الكمى ، والاستدلال المجرد فى مقابل الاستدلال البصرى ، والذاكرة قصيرة المدى ، وبالاضافة الى ذلك يمكن أن يعطى المفحوص درجة كلية مركبة فى المقياس ٠

وعلى الرغم من أن كل درجة من هذه الدرجات تعكس جزئيا أحد عوامل القدرة العقلية العامة المشتركة فى جميع القدرات ، وعلى الرغم ايضا من الدرجة الكلية المركبة من هذه الدرجات الأربع تعطى تقديرا ثابتا للعامل العام ، فان كل درجة منها لها معناها السيكولوجى الخاص والذى له أهميته فى ذاته ٠ وقد حسبت المعايير لكل درجة من الدرجات الأربع وللدرجة الكلية ٠ وقد نظم المقياس تبعا لمستويات تبدأ من الطفل الأقل من المتوسط من سن سنتين وحتى مستوى الراشد المتفوق ٠

تحليل أسئلة المقياس :

لقد استغرق تعديل عام ١٩٣٧ لمقياس ستانفرد - بينيه أكثر من عشر

سنوات من البحث العلمى المنظم حول تجميع الأسئلة الجيدة وتجريبها على جماعات صغيرة العدد من المفحوصين الذين طبق عليهم مقياس عام ١٩١٦٠ وكانت الخطوة الأساسية فى ذلك هى تطبيق الصورتين المبدئيتين فى ذلك الوقت على عينة التقنين التى أشرنا اليها (٣١٨٤ مفحوصا) ومنها تم الانتقاء النهائى للاسئلة وتوزيعها على الاعمار وقد أعطيت لكل مفحوص صورتا الاختبار بفاصل زمنى يتراوح بين يوم واحد وأسبوع واحد ٠ وفى هذه الخطوة كان محك تحليل الأسئلة لتحديد صدقها هو العمـــر الزمنى والدرجة الكلية المركبة من الصورتين المبدئيتين ، ويرتبط المحك الثانى بما اسميناه الاتساق الداخلى حتى يتحقق فى المقياس درجة من التجانس ٠ واستخدمت ثلاثة مقاييس فى تحليل كل سؤال فى المقياس هى :

١ ـ منحنيات النسب المئوية للمفحوصين الذين يجتازون الســؤال بنجاح فى الأعمار الزمنية المتتابعة ٠

٢ ـ منحنيات النسب المئوية للمفحوصين الذين يجتازون الســؤال بنجاح فى الفئات المتتابعة للدرجة الكلية فى المقياس (وبالطبع لكل صورة من صورتى المقياس منحنياتها الخاصة) ٠

٣ ـ معامل الارتباط الثنائى بين كل سؤال والدرجة الكلية فى المقياس٠

وتتفق الطريقة الأولى مع أسلوب بينيه الذى كان يستخدمه فى تحليل المفردات وانتقاء الأسئلة التى تصلح للمستويات العمرية المختلفة ٠ وفى هذا يجب ان نشير الى أن ترمان استخدم للحكم على صلاحية الأسئلة نسبا مئوية فى الأعمار الصغيرة أعلى منها فى الأعمار الكبيرة ٠ ومن ذلك مثلا انه كان يعتبر السؤال مناسبا لسن ٣ سنوات اذا اجاب عليه ٧٣٪ من أطفال هذا السن اجابة صحيحة ، أما فى حالة الأطفال من سن ١٠ سنوات فكانت هذه النسبة ٥٩٪ ٠ ويشير ماكنمر (٦٩) الى أن هذا الاختلاف فى النسب المئوية هو من المطالب الضرورية فى مقاييس العمـــر ٠ فمن الملاحظ أن الانحراف المعيارى للعمر العقلى يجب أن يتزايد مع العمر اذا كان لابد لنسبة الذكاء أن تظل ثابتة ٠ وعلى ذلك فاذا كان التباين فى العمـــر العقلى فى الأعمار العليا اكبر منه فى الأعمار الدنيا فان ذلك يعنى ضرورة وجود انتشار اكبر فى أداء المفحوصين الكبار فى مستويات الأعمار المتجاورة ٠ وعلى ذلك

فمع الزيادة فى العمر يقل عدد المفحوصين الذين يجتازون بنجاح السؤال المطابق لمستواهم العمرى ٠

ويمكن أن نوضح هذه المسألة بمثال بسيط لنفرض انه بين الاطفال من سن ٥ سنوات يصل ٤٠٪ منهم الى العمر العقلى ٥ ، بينما يصل ٣٠٪ الى اعمار اعلى ، و ٣٠٪ آخرون الى اعمار أقل من هذا المستوى ٠ وحيث أن اى مفحوص عمره العقلى ٥ سنوات أو اكبر يجب ان يجتـــاز من الوجهة النظرية السؤال الخاص بسن ٥ سنوات فان ٧٠٪ (٤٠٪ + ٣٠٪) من الأطفال من فئة سن ٥ سنوات تتوقع نجاحهم فى أسئلة سن ٥ ســـنوات ٠ ولنفرض من ناحية أخرى أنه من بين الأطفال من سن ١٠ سنوات لا يصل الى العمر العقلى ١٠ سنوات الا ٢٠٪ ، بينما يصل الى اعمار أعلى من ذلك ٤٠٪ منهم ، والى أعمار أقل من ذلك ٤٠٪ آخرون ٠ وبالطبع لا يمكن الحصول على مثل هذا التوزيع الا اذا كان التباين فى الأعمار العقلية لأطفال سن ١٠ سنوات أكبر منه فى سن ٥ سنوات ٠ ويعنى ذلك أنه من بين الاطفال من سن ١٠ سنوات تصبح النسبة المتوقعة لاجتياز الأسئلة المخصصة لهذا السن هى (٢٠٪ + ٤٠٪ = ٦٠٪) وفى هذا المثال تصبح النسبة المئوية التى يمكن استخدامها محكا لانتقاء الأسئلة فى سن ٥ سنوات هى ٧٠٪ ، وفى سن ١٠ سنوات هى ٦٠٪ ٠

وفى الممارسة الواقعية لانتقاء اسئلة مقياس ستانفرد ـ بينيه نجد أن المحك يتناقص من ٧٧٪ عند مستوى سن عامين الى ما هو أقل قليلا من ٥٠٪ فى مستوى الرشد العادى ٠ وتقل النسب عن هذا فى انتقاء أســئلة مستويات الرشد المتفوق وذلك للوصول الى سقف مناسب للاختبار ٠ وعند الانتقاء النهائى للمفردات لا يقتصر الاهتمام على تمايز العمر والاتســاق الداخلى فحسب ، وانما امتد ليشمل أيضا مدى تناقص الفروق بين الجنسين وتوزانها فى النسب المئوية للناجحين فى مختلف المســتويات العمرية ٠ واستبعدت بالفعل المفردات التى اختلف فيها اداء احد الجنسين عن الآخر بنسب لها دلالة احصائية : على اساس أن مثل هذه المفردات قد تدل على تباين اصطناعى أو غير مرتبط بما يقيسه الاختبار نتيجة للخبرات السابقة للجنسين ٠

وفى اعداد طبعة عام ١٩٦٠ من المقياس اختبرت الأسئلة من الصورتين ل، م على أساس أداء ٤٤٩٨ مفحوصا يمتد اعمارهم بين ٢½ سنة ، ١٨ سنة اجابوا على احدى صورتى المقياس او كليهما فى الفترة من ١٩٥٠ وحتى ١٩٥٤ . كما تضمنت العينة مجموعتين طبقيتين اشتملتا على ١٠٠ طفل من سن ٦ سنوات صنفت حسب مهنة الوالد ، ١٠٠ مراهق من سن ١٥ سنة صنفت حسب مهنة الوالد وتوزيع المستوى التعليمى .

ويعنى هذا أن طبعة ١٩٦٠ لم تتضمن اعادة لتقنين المقياس أو اعادة لبناء معاييره . واستخدمت العينات الجديدة لمراجعة التغيرات التى طرأت على صعوبة الأسئلة فقط بعد ما تحددت مستويات الصعوبة بايجاد النسب المئوية للاطفال الذين يجتازون السؤال فى أعمار عقلية متتابعة فى صورتى عام ١٩٣٧ وهكذا فانه بالنسبة لتحليل المفردات تم تجميع المفحوصين تبعا للعمر العقلى الذى تحدده طبعة عام ١٩٣٧ وليس تبعا للعمر الزمنى وعلى هذا فان الأعمار العقلية ونسب الذكاء فى طبعة ١٩٦٠ ظلت على النحو الذى كانت عليه فى طبعة ١٩٣٧ وفى ضوء أداء عينات تقنين هذه الطبعة . صحيح أنه روجعت بعض الأسئلة للتحقق من احتمال وجود فروق جغـــرافية أو اقتصادية اجتماعية وذلك من خلال المقارنة بين المجموعات الفرعية التى تكون العينة الكلية . الا ان ما يجب أن ننبه اليه أن المقياس لم يتضمن أى مـــواد جديدة ، وانما عدلت بعض الرسوم الخاصة بالأشياء الشائعة مســــايرة للعصر . وفى هذه الأحوال اختبرت الأسئلة المعدلة قبليا على عينات صغيرة قبل تضمينها فى المقياس . وقد ظلت هذه الاستراتيجية فى انتقاء المفردات ثابتة فى طبعتى ١٩٧٢ ، ١٩٨٥ .

وصف المقياس :

يتألف مقياس ستانفرد - بينيه من صندوق به مجموعة من اللعب المقننة تستخدم مع الأعمار الدنيا ، وكتيبين من البطاقات المطبوعة ، وكراســـة لتسجيل الاجابات ، وكراسة التعليمات ، وكراسة معايير التصحيح .

وقد رتبت أسئلة المقياس وجمعت تبعا لمستويات الأعمار التى تمتد من سن عامين حتى سن الرشد المتفوق . وتتحدد مستويات المقياس من سن ٢

حتى سن ٥ سنوات فى فئات نصف سنوية على النحو الآتى : ٢ ، ٢ ـ ٦ ، ٣ ، ٣ ـ ٦ ، ٤ ، ٤ ـ ٦ ، ٥ (مع ملاحظة أن الرمز ٢ ـ ٦ ، ٣ ـ ٦ ، ٤ ـ ٦ ، يدل على سن عامين و ٦ شهور ، ٣ سنوات و ٦ شهور ، ٤ سنوات و ٦ شهور على التوالى) وذلك بسبب التقدم العقلى السريع فى الأعمار الدنيا · أما الفترات من سن ٥ سنوات حتى ١٤ سنة فتتحدد مستويات الأعمار فيها فى فئات سنوية ، وتدل المستويات الأخرى فى المقياس على الراشد المتوسط ، والراشد المتفوق (١ ، ٢ ، ٣) · وكان عدد اسئلة كل مستوى عمرى ٦ ، باستثناء مستوى الراشد المتوسط الذى يبلغ عدد اسئلته ٨ ·

وكانت الأسئلة فى كل مستوى عمرى من مستوى صعوبة موحدة تقريبا ، وترتب بصرف النظر عن أى فروق طفيفة فى الصعوبة · وقد أعد سؤال احتياطى فى كل مستوى عمرى معادل فى صعوبته لأسئلة العمر · ويستخدم هذا السؤال الاحتياطى اذا تطلب الأمر استبعاد احد الاسئلة العادية بسبب بعض الظروف الخاصة تجعله غير ملائم للمفحوص ، أو بسبب عدم الالتزام فى التطبيق بحرفية التعليمات بما يتعارض مع تقنين المقياس وهكذا يبلغ العدد الكلى لاختبارات المقياس ١٤٢ اختبارا (١٢٢ اختبارا اصليا ، ٢٠ اختبارا احتياطيا) ·

ومعنى هذا أنه ابتداء من طبعة ١٩٦٠ من المقياس نجد سؤالا احتياطيا لكل مستوى عمرى ، وهذا مالم يتوافر فى طبعة ١٩٣٧ التى نقلت الى اللغة العربية حيث اقتصر وجود الأسئلة الاحتياطية على المستويات العمرية من سن ٢ حتى سن ٥ سنوات فقط ·

وقد اختير ٤ أسئلة من كل مستوى عمرى على أساس الصدق والتمثيل ليتألف منها المقياس المختصر الذى يستخدم حين لا يسمح الوقت باعطاء المقياس الكلى ، وقد وضعت امام هذه الاختبارات علامة (*) · وتؤكد المقارنات بين نسب الذكاء كما تتحدد بالمقياس الكلى والمقياس المختصر تطابقا كبيرا بينهما حيث يصل معامل الارتباط بينهما الى مستوى معامل ثبات المقياس الكلى (٥٦ ، ٥٧) ·

تطبيق المقيـــاس :

يتطلب هذا المقياس - كغيره من اختبارات الذكاء الفردية - فاحصا على درجة كبيرة من المهارة والتدريب لان تطبيق هذا المقياس وتصحيحة على درجة كبيرة من التعقد ويستلزم أن تتوافر فى الفاحص الفة شديدة به وخبرة طويلة فى التطبيق · لأن تردد الفاحص اثناء التطبيق قد يؤدى الى عجز عن تكوين علاقة بينه وبين المفحوص ، بل وفقدان ثقة المفحوص به · بل ان أدنى تغيير فى صياغة الأسئلة قد يؤدى الى تغيير فى مستوى صعوبتها كما تتحدد فى عينة التقنين · بالاضافة الى ما يتطلبه المقياس من تصحيح فورى لاختباراته عقب اعطائها مباشرة ، لأن تتابع اعطاء الاختبارات يعتمد على أداء المفحوص فى المستويات السابقة ·

ويرى كثير من علماء النفس الاكلينيكيين أن مقياس ستانفرد - بينيه ليس مقياسا مقننا للذكاء فحسب·وانما هو كذلك نوع من المقابلة الاكلينيكية، لأنه يهيىء فرصة للتفاعل بين الفاحص والمفحوص ، ويهيىء مصادر أخرى للدلالات الاكلينيكية لدى الأخصائى الاكلينيكى المتمـــرس ، بل ان مقياس ستانفرد - بينيه ، أكثر من غيره من الاختبارات الفردية ، يهيىء الفرصـــة للفاحص أن يلاحظ الطرق التى يلجأ اليها المفحوص فى العمل ، وأساليب حله للمشكلات ، وغير ذلك من الأساليب الكيفية للاداء ، بل ان الفاحص يمكنه أن يحكم من خلال الموقف الاختبارى الذى يستخدم فيه هذا المقيـاس على بعض خصائص الشخصية مثل مستوى النشاط والثقة بالنفس والمثابرة والقدرة على التركيز · وبالطبع فان أى ملاحظات كيفية يمكن تسجيلها أثناء موقف الاختبار يجب تحديدها فى ذاتها ، ولا يجب تفسيرها على نفس النحو الذى تفسر به درجات الاختبار الموضوعى · وتعتمد قيمة هـــذه الملاحظات الكيفية الى حد كبير على مهارة الفاحص وخبرته واعداده السيكولوجى العميق ، بالاضافة الى وعيه الشديد بمزالق وحدود هذا النوع من الملاحظات المباشرة · وقد استطاع موريايترى (٧٠) أن يســـجل انماط الملاحظـــات الاكلينيكية فى موقف الاختبار الفردى ، على أساس أن جلسة الاختبار هى لدراسة سـلوك الطفل أثناء مواجهة موقف تسود فيه عوامل التحـــدى والاستثارة والصعوبة والاحباط ·

وفى اعطاء مقياس ستانفرد ـ بينيه لا يحاول المفحوص جميع أسئلة المقياس ، وانما يختبر المفحوص فقط حول مدى المستويات العمرية الملائمة لمستواه العقلى • ولا يتطلب المقياس أكثر من ٣٠ ـ ٤٠ دقيقة للاطفـــال الصغار ، ولا يتطلب أكثر من ساعة ونصف للراشدين • والاجراء المقنن فى تطبيق المقياس هو أن نبدأ الاختبار من مستوى أدنى قليلا من العمر العقلى المتوقع للمفحوص وعلى ذلك فان الاختبارات الأولى يجب ان تكون سهلة بالقدر الذى يستثير فى المفحوص الثقة بالنفس ، بشرط الا تكون سهلة جدا بحيث تؤدى بالمفحوص الى الاستهانة بالاختبار أو الملل • والاجراء الأكثر موضوعية فى هذا الصدد هو أن نبدأ باعطاء المفحوص أسئلة العمر الأدنى مباشرة من عمره الزمنى ، فاذا فشل فى أى سؤال من أسئلة هذا المستوى يعطى أسئلة المستوى الأدنى منه ، وهكذا يستمر الفاحص حتى يصل الى العمر الذى يستطيع المفحوص أن يجيب على جميع اسئلته ، ويصبح هذا المستوى هو « العمر القاعدى » basal age وبعد ذلك نعطيه أسئلة عمره الزمنى فان أجاب عليها جميعا أو على بعضها تعطيه أسئلة الأعمار الاعلى منه تدريجيا حتى نصل الى العمر الذى يفشل المفحوص فى الاجابة على جميع أسئلته ويسمى هذا المستوى « الحد الأعلى للعمر » أو سقف الاختبار ceiling age • وحين يصل الفاحص الى هذا المستوى يتوقف عن اعطاء المقياس •

تصحيح المقياس :

تصحح جميع أسئلة مقياس ستانفورد ـ بينيه على أساس النجاح أو الفشل (واحد أو صفر) • وتحدد كراسة التعليمات الأداء المطلوب « للنجاح » فى كل سؤال •وقد تظهر نفس الأسئلة فى مستويات عمرية مختلفة ولكنها تصحح بمستويات مختلفة للنجاح • وفى هذه الحالة تطبق الأسئلة مـرة واحدة ثم يحدد أداء المفحوص المستوى العمرى الذى يستحقه • ومن ذلك مثلا أن سؤال المفردات يطبق ابتداء من المستوى العمرى ٦ سنوات حتى المستوى الثالث للراشد المتفوق اعتمادا على عدد الكلمات التى يستطيع المفحوص تحديدها وتعريفها •

والواقع أن الأسئلة التى ينجح فيها المفحوص أو يفشل تظهر مقدار من التشتت فى المستويات العمرية المتتابعة • فنحن لا نجد مفحوصا يجيب على جميع أسئلة عمره العقلى أو أدنى منه ويفشل فى جميع الاسئلة الأعلى من هذا المستوى ، وانما نجد فى العادة ان الأسئلة التى ينجح فيها تنتشر فى مستويات عمرية عديدة يحددها العمر القاعدى من ناحية والحد الأعلى للعمر (سقف الاختبار) من ناحية أخرى • ويحسب العمر العقلى فى هذه الحالة بالعمر القاعدى مضافا اليه أى شهور عقلية لأى اختبار اجتازه من مستويات أعلى من عمره القاعدى تبعا للاوزان المبينة فى الجدول رقم (٨)•

جدول رقم (٨) يبين أوزان أسئلة مقياس ستانفرد ـ بينيه فى الأعمار المختلفة

العمر	عدد الأسئلة	وزن السؤال بالشهر	مجموع التقدير الوزنى لكل مستوى بالشهر
٢	٦	١	٦
٢ ـ ٦	٦	١	٦
٣	٦	١	٦
٣ ـ ٦	٦	١	٦
٤	٦	١	٦
٤ ـ ٦	٦	١	٦
٥	٦	١	٦
٦	٦	٢	١٢
٧	٦	٢	١٢
٨	٦	٢	١٢
٩	٦	٢	١٢
١٠	٦	٢	١٢
١١	٦	٢	١٢
١٢	٦	٢	١٢
١٣	٦	٢	١٢
١٤	٦	٢	١٢
الراشد المتوسط	٨	٢	١٦
الراشد المتفوق (١)	٦	٤	٢٤
الراشد المتفوق (٢)	٦	٥	٣٠
الراشد المتفوق (٣)	٦	٦	٣٦
المجموع	١٢٢	المجموع	٢٥٦
	+		+
عدد الأسلة الاحتياطية	٢٠	العمر الأدنى لبدء الاختبار	١٨
المجموع الكلى	١٤٢	شهرا	٢٧٤

شهور سنة

اى ١٠ ٢٢

ومن الوجهة النظرية فان أعلى عمر عقلى يمكن الوصول اليه بمقياس ستانفرد ـ بينيه ـ هو ٢٢ سنة و ١٠ شهور ٠ والواقع ان هذه الدرجة لاتدل على عمر عقلى حقيقى ، وانما هى درجة عددية تدل على مقدار الزيادة عن الأداء المتوسط الذى يصدر عن الراشد العادى ٠ وهى بالتأكيد لا تطابق ما يصل اليه الشخص المتوسط من عمر زمنى مقداره ٢٢ سنة و ١٠ شهور، لأن هذا الشخص العادى يحصل على عمر عقلى مقداره ١٦ سنة و ٨ شهور كما تحدده معايير ١٩٧٢ (*) ٠ أما بالنسبة لأى راشد يزيد عمره عن ١٨ سنة فان أى عمر عقلى مقداره ١٦ سنة ، ٨ شهور يحصل على نسبة الذكاء ١٠٠٠

ويجب أن نؤكد هنا أنه بالنسبة للمفحوصين ابتداء من مرحلة المراهقة (١٣ +) لا يعطى العمر العقلى فى هذا المقياس نفس الدلالة التى يعطيها فى الاعمار الأدنى ، لأنه بعد سن ١٣ يقل متوسط الأعمار العقلية عن متوسط الأعمار الزمنية ، ولذلك فان مقياس ستانفرد ـ بينيه لا يلائم الراشـــدين وبخاصة العاديين والمتفوقين منهم ٠ ورغم وجود ٣ مستويات للراشد المتفوق فيه فان « سقف » الاختبار لا يلائم معظم الراشدين المتفوقين أو حتى الراشدين العاديين كما تؤكد دراسات كيندى وزملائه (٦٧) فمن المستحيل ان نجـد فى المقياس سقفا مناسبا لهؤلاء ، اى مستوى عمرى يعجز المفحوصـــون فى الاجابة عن جميع أسئلته وبالاضافة الى ذلك فان مقياس ستانفرد ـ بينيه اكثر جاذبية للصغار منه للمراهقين والراشدين من حيث محتواه ٠

ومن أهم الاضافات الى طبعة عام ١٩٦٠ احلال نسبة الذكاء الانحرافية محل نسبة الذكاء العادية التى استخدمت فى المقياس منذ ظهوره عام ١٩١٩ ٠ وهذه النسبة الجديدة عبارة عن درجة معيارية معـــدلة متوســـطها ١٠٠ وانحرافها المعيارى ١٦ ٠والميزة الأساسية فى هذا النوع من المعايير أنها تعطى درجات قابلة للمقارنة فى مختلف الأعمار ٠ وبهذا امكن التغلب على مشكلة عدم استقرار نسبة الذكاء العادية ٠ فعلى الرغم من الجهود التى بذلت عام ١٩٣٧ للحصول على تشتتات ثابته لنســـبة الذكاء فى مختلف الأعمار ، فاننا نلاحظ أن هذا المقياس تذبذبت فيه الانحـرافات المعيارية

(*) كان هذا العمر ١٦ سنة فى طبعة ١٩٣٧ ٠

لمختلف الأعمار من انحراف معيارى ١٣ (فى سن ٦) الى انحراف معيارى ٢١ (فى سن ٢ ـ ٦) ومعنى هذا أن نسبة الذكاء ١١٣ فى سن ٦ تقابل نسبة الذكاء ١٢١ فى سن ٢ ـ ٦ ٠ وقد وضع مكنمار جداول خاصة بتصحيح آثار هذه الاختلافات فى طبعة عام ١٩٣٧ (فى ٧٨) وتتلخص فكرة التصحيح فى أنه اذا كان العمر العقلى للمفحوص بين سن ١٣ ، ١٦ فانه يجب تعديل العمر الزمنى الذى يؤخذ فى الاعتبار عند حساب نسبة الذكاء ، وذلك بأن نضيف ثلثى الشهور الاضافية التى عاشها المفحوص بعد سن ١٣. فمثلا اذا كان العمر الزمنى الحقيقى ١٤ سنة فانه يحسب ١٣ ـ ٨ ، واذا كان ١٥ سنة يحسب ١٤ ـ ٤ ٠

وقد أمكن التغلب على كل هذه الصعوبات فى طبعة ١٩٦٠ باستخدام نسبة الذكاء الانحرافية والتى تعنى استخدام انحراف معيارى موحد لجميع الأعمار ٠ وقد أعد بينو (٧٢) جداول تستخرج منها مباشرة نسبة الذكاء الانحرافية من كل من العمر العقلى والعمر الزمنى بالسنوات والشهور ٠ وقد نشرت هذه الجداول فى كراسة تعليمات المقياس الصادرة عام ١٩٧٣ (٨٠) ٠

وتعتمد فكرة التحويل فى جداول بينو على اعتبار نسبة الذكاء التقليدية هى الدرجة الخام للمفحوص تحول الى نسبة ذكاء انحرافية باستخدام متوسطات وانحرافات معيارية تجريبية لنسب الذكاء التقليدية لكل فئة من الفئات العمرية ، ثم يستخدم الانحراف المعيارى الجديد ١٦ والمتوسط الجديد ١٠٠ ، اى أن معادلة نسبة الذكاء الانحرافية فى مقياس ستانفرد ـ بينيه هى:

نسبة الذكاء الانحرافية =

$$\left[١٦ \times \left(\frac{\text{نسبة الذكاء التقليدية ـ المتوسط التجريبى}}{\text{الانحراف المعيارى التجريبى}} \right) \right] + ١٠٠$$

أو

$$\left[\left(\text{نسبة الذكاء التقليدية ـ المتوسط التجريبى} \right) \times \frac{١٦}{\text{الانحراف المعيارى}} \right] + ١٠٠$$

وقام بينو بقسمة ١٦ على الانحراف المعيارى التجريبى لكل فئة عمرية فحصل على ما اسماه بالمقدار الثابت ك K واعد جدولا بهذه الثوابت لفئات عمرية مداها شهر ٠ والجدول رقم (٩) يوضح هـــذه الثوابت لفئات العمر المختلفة الا ان ما يجب أن نشير اليه هو أن حســاب العمر العقلى لا يزال يتم بالطريقة المستخدمة منذ عام ١٩٣٧ ٠ اى ان التعديلات ادخلت على مفهوم نسبة الذكاء التقليدية دون أن تمس مفهوم العمر العقلى ذاته ٠

جدول (٩) يبين الثوابت اللازمة لتحويل نسبة الذكاء التقليدية الى نسبة ذكاء انحرافية ٠

العمر الزمنى	المتوسط التجريبى	ك	العمر الزمنى	المتوسط التجريبى	ك
٢	١٠٢	٩٩ر			
٢ — ٦	١٠٣	٧ر	٨	١٠٢	١ر٠٢
٣	١٠٤	٨٨ر	٩	١٠٣	٩٨ر
٣ — ٦	١٠٢	٩٣ر	١٠	١٠٣	٩٤ر
٤	١٠٢	٩٨ر	١١	١٠٢	٩١ر
٤ — ٦	١٠١	١ر٠٤	١٢	١٠٢	٨٨ر
٥	١٠٠	١ر٠٩	١٣	١٠١	٨٩ر
٦	١٠٠	١ر١١	١٤	١٠١	٩٠ر
٧	١٠١	١ر٠٧	الراشـــد	١٠٥	٩٠ر

ويوجد تعديل آخر اضافته طبعة ١٩٦٠ واكدته طبعة ١٩٧٣ ٠ واعتمد على النتائج التى تؤكد أن التحسن فى الأداء فى المقياس يستمر الى ســـن ١٨ سنة بدلا من التـــوقف عنــد سن ١٦ كمـــا كان الحـــال فى طبعـــة ١٩٣٧ ٠ فقد بنيت البحوث الطولية فى النمو العقلى التى اجريت خلال الربع قرن بين طبعتى المقياس أن القدرات العقلية كما تقاس باختبـــارات الذكاء

تستمر فى التحسن والارتقاء أبعد مما كان مفترضا (٢٤) · واكثر الأدلة مباشرة على زيادة التحسن فى مقياس ستانفرد - بينيه بعد سن ١٦ قدمته بحوث برادوى وطومسون وكرافنر (٨٢) ، بعد انقضاء ١٠ سنوات ثم بعد انقضاء ٢٥ سنة ، فلوحظ أن هؤلاء المفحوصين حين اختبـــروا أول الأمر وكانت اعمارهم الزمنية ٢ - ٥½ سنة ثم اختبروا بعد ١٠ ســـنوات ، أن متوسطهم فى نسبة الذكاء لم يختلف فى الحالتين عما حددته معايير المقياس ولكن هذا المتوسط اظهر ارتفاعا له دلالة احصائية بمقدار ١١٫٣ نقطـة بين الاختبار الثانى والثالث (أى بعد انقضاء ١٥ سنة من الاختبار الثانى و ٢٥ سنة من الاختبار الأول) · وهذه الزيادة نتجت عن ان المفحوصين استمروا فى التحسن بعد سن ١٦ سنة ، بينما نجد ان حساب نســـب الذكاء الذى اعتمدت عليه طبعة ١٩٣٧ افترض توقف النمو العقلى عند هذا السن ·

وعند المقارنة بين طبعات ١٩٣٧ ، ١٩٧٢ ، ١٩٨٥ نجد ان المعايير فى الطبعات الاحدث تعتمد على عينات اكثر تمثيلا ، كما انها اكثر حداثة وتعكس أى تغيرات فى الظروف الحضارية والثقافية يمـــكن ان تؤثر فى الأداء الاختبارى · ومن الطريف أن نذكر ان المعايير الأحدث تظهر بعض التحسن فى أداء الاختبار فى جميع الأعمار ، ويكون هذا التحسن اكثر جوهرية فى اعمار ما قبل المدرسة يصل فى متوسطه الى حوالى عشر نقـــاط · ويرجع مؤلفو الاختبار هذا التحسن الى اثر وسائل الاعلام من ناحية والى ارتفاع المستوى الثقافى والتعليمى للوالدين ·

طريقة حساب العمر العقلى ونسبة الذكاء التقليدية :

يمكن توضيح الأسلوب الذى يستخدم فى الحصول على العمر العقلى ونسبة الذكاء التقليدية من مقياس ستانفرد - بينيه بالمثال الموضح فى الجدول رقـــم (١٠) ·

جدول رقم (١٠) يوضح طريقة حساب العمر العقلى ونسبة الذكاء التقليدية

العمر	عدد الأسئلة التى أجاب عليها المفحوص	وزن السؤال بالشهر .	المجموع الكلى للأوزان (شهر)	المجموع الكلى للأوزان (سنة)
طفل عمره الزمنى ٦ سنوات ، ٤ شهور			شهر	سنة
٤	٦	العمر القاعدى	٢	٤
٤ - ٦	٥	١	٥	—
٥	٣	١	٣	—
٦	٣	٢	٦	—
٧	٢	٢	٤	—
٨	١	٢	٢	—
٩	صفر	سقف الإختبار	—	—
			—	٤

		شهر	سنة	
العمر العقلى	=	٨	٥	نسبـــة الذكاء
العمر الزمنى	=	٤	٦	التقليدية = ٨٩

وبالطبع تحول نسبة الذكاء التقليدية ٨٩ الى نسبة ذكاء انحرافية باستخدام جداول بينو المشار اليها .

تفسير معايير المقياس :

من أهم مزايا مقياس ستانفرد ـ بينيه ما يتوافر حـــوله من بيانات امبريقية وخبرة كلينيكية تراكمت طوال ثمانين عاما منذ مقياس بينيه المبكر. وقد أصبح هذا المقياس عند كثير من علماء النفس الكلينيكيين والتربويين وغيرهم من المهمتين بالتقويم النفسى مرادفا للذكاء . وأصبح تحديد مستويات الذكاء يعتمد فى جوهره على نسب الذكاء كما يحددها المقياس ، رغم ما

فى هذا الرأى من غلو لا يتسع المقام لمناقشته • والجدول رقم (١١) يوضح التوزيع والتصنيف النظريين لنسب الذكاء الانحـــرافية لطبعتى ١٩٦٠ ، ١٩٧٢ •

جدول رقم (١١) يبين التوزيع والتصنيف النظرى لنسب الذكاء الانحرافية لمقياس ستانفرد ـ بينيه

الدرجة المعيارية	نسب الذكاء الانحرافية	التصنيف	النسب المئوية للحالات
ـ ٣	٥٢	التخلف العقلى (الضعف العقلى)	٢ر١ ٪
ـ ٢	٦٤ ـ ٦٨ ـ ٧٦	الحالات الهامشية (قد تصنف أغبياء أو ضعاف عقول	٥ر٥٪
ـ ١	٨٤ ـ ٨٨	الغباء (بطء التعلم)	١٦٪
صفر	١٠٠ ـ ١١٢	المتوسط (أو العادى)	٧ر٥٤٪
+ ١	١١٦ ـ ١٢٤	المتفوقون	١٦ ٪
+ ٢ ، + ٣	١٣٢ ـ ١٤٨	المتفوقون جداً	٥ر٦٪
	أعلى من ذلك	العبقرية أو شبه العبقرية	١ر٪

وقد كان من المعتاد لعدة سنوات استخدام نسبة الذكاء التقليدية ٧٠ باعتبارها نقطة قطع تقريبية للتخلف العقلى او الضعف العقلى المحتمل ، وتتطابق هذه الفئة مع النسبة ١٪ أو ٢٪ من الحد الأدنى للتوزيع • أما الجدول فيوضح لنا أن نقطة القطع هذه أصبحت عند الدرجة المعيـــارية (ـ ٢٨ر٢) والتى تتطابلا مع نسبة الذكاء الانحرافية ٦٤ والتى يقع أدنى منها ٢ر١٪ من الحالات •

وعموما فان الفضل يعود الى هذا المقياس فى استخدام التقسيم الثلاثى الشهير للضعف العقلي : المأفونون والبلهاء والمعتوهون ، وهى الفاظ لم تستخدم فى الوقت الحاضر وحل محلها ما يلى :

١ _ الحالات القابلة للتعلم educable وتمتد نسب ذكائها من ٥٠ ، ٧٥ ويمكنها أن تتعلم حتى الصف الثالث فى المرحلة الابتدائية ولا يتجاوز حدها الاقصى فى التعلم نهاية المرحلة الابتدائية اذا تعلمت موقف تعليمى خاص ·

٢ _ الحالات القابلة للتدريب trainable وتمتد نسب ذكائها بين ٢٥ ، ٥٠ ويمكنها أن تتعلم رعاية الذات والتوافق الاجتماعى فى بيئة حماية خاصة ·

٣ _ الحالات التى تحتاج الى الايداع فى المؤسسات custodial وتقل نسب ذكائها عن ٢٥ وتتطلب رعاية وحماية خاصة ·

والتصنيف الراهن للضعف العقلى يقسم هذه الحالات الى اربع فئات يوضحها الجدول رقم (١٢) ·

جدول رقم (١٢) يبين مستويات الضعف العقلى فى صورة نسب ذكاء انحرافية

المستوى	نقاط القطع (وحدات ع)	مدى نسب الذكاء والانحراف (ع = ١٦)	النسب المئوية للحالات
الضعف العقلى الخفيف	_ ٢	٥٢ _ ٦٨	١٤ر٣
الضعف العقلى المتوسط	_ ٣	٣٦ _ ٥١	١٣ر
الضعف العقلى الشديد	_ ٤	٢٠ _ ٣٥	٠٠٣ر
الضعف العقلى الخطير	_ ٥	١٩ فأدنى	٠٠٠٠٣ر

والميزة الجوهرية فى هذا التصنيف انه يصلح للتحويل الى درجات معيارية أو نسب ذكاء انحرافية فى اى مقياس للذكاء ٠

والواقع أن شيوع تصنيف مستويات نسب الذكاء - رغم فائدته فى تقنين تفسير الأداء الاختبارى - يتضمن بعض المخاطر ٠ فهو كغيره من صور التصنيف السلوكى يجب أن يستخدم بمرونة كافية ، كما يجب الا يستبعد البيانات الأخرى التى تتوافر عن المفحوص ٠ وبالطبع لا يوجد حد فاصل تماما بين فئات التصنيف ٠ ومن ذلك مثلا لا يوجد خط يميز تماما بين التخلف العقلى والحالات الهامشية ، أو بين التفوق والتفوق الشديد ٠ فقد لوحظ مثلا أن بعض المفحوصين من ذوى نسب الذكاء ٦٠ يتكيفون تكيفا عاديا لمطالب الحياة اليومية ، بينما نجد بعض الأشخاص من ذوى نسب الذكاء ١٠٠ يتطلبون الايداع فى المؤسسات ، كما أن بعض الاشخاص من ذوى نسب الذكاء ١٦٠ قد لا يعيشون حياة متميزة ، بينما نجد البعض الآخر من ذوى نسب الذكاء الأقرب الى ١٠٠ يسهمون بعض الاسهامات الممتازة ٠ ومعنى ذلك أن القرارات التى تتحدد حول تشخيص الضعف العقلى أو التفوق العقلى يجب الا تعتمد على نسب الذكاء وحدها ، وانما يجب فى حالة الضعف العقلى والقرارات المتصلة به كالايداع فى المؤسسات أو الالتحاق بمدارس التربية الفكرية الاهتمام بمتغيرات أخرى مثل النضج الاجتماعى والتوافق الانفعالى والظروف الصحية والجسمية والأحوال الاسرية والظروف المدرسية وغير ذلك من العوامل التى تؤثر فى الآداء العقلى وفى جميع الحالات يجب الاهتمام بكل من النشاط العقلى للفرد ونضجه الاجتماعى وسلوكه التوافقى جميعا وفى وقت واحد ٠ وبالمثل فان نسب الذكاء العالية لا ترادف العبقرية ، لأن الانجاز الرفيع قد يتطلب بالاضافة الى ذلك الابتكار والمواهب الخاصة والمثابرة والاهتمام بالهدف وغيرها من العوامل العقلية والانفعالية والدافعية (٣٢) ٠

ثبات المقياس :

اجريت دراسات الثبات على طبعة ١٩٣٧ من المقياس باستخدام طريقة الصور المتكافئة على فترات زمنية مقدارها اسبوع أو اقل (للتحقق من مصدرى تباين الخطا : الاستقرار الزمنى وتكافؤ المحتوى) ٠ وقد لوحظ

أن المقياس يميل الى أن يكون أكثر ثباتا فى الأعمار الكبيرة منه فى الأعمار الصغيرة ، وفى مستويات نسب الذكاء الدنيا عنه فى نسب الذكاء العليا ٠

والواقع أن زيادة الثبات مع الزيادة فى العمـــر هى من خصائص الاختبارات النفسية بوجه عام ، وتنتج جزئيا عن التحكم الأفضل فى ظروف اعطاء الاختبار فى حالة المفحوصين الكبار السن (وبخاصة حين نقارنهم بأطفال ما قبل المدرسة) ٠ بالاضافة الى العوامل الأخرى المسهمة فى زيادة معاملات الثبات ومنها بطء معدل النمو مع التقدم فى العمر ٠

اما عن ارتفاع معاملات الثبات فى حالة نسب الذكاء الدنيا فى ى عمر زمنى فترتبط فيما يبدو بخصائص مقياس ســتانفرد ـ بينيه على وجـــه الخصوص ٠ فقد أشرنا الى أن الأوزان النسبية لأســئلة الأعمار المختلفة تختلف ، فهى فى الأعمار الدنيا مقدارها شهر ، وفى الاعمار المتوســطة شهران ، وفى المستويات العليا قد تكون ٤ شهور أو ٥ شهور أو ٦ شهور٠ ومعنى ذلك أن طريقة وزن الدرجات تميل الى « تضخيم » اخطاء القياس فى المستويات العليا ، لأن النجاح أو الفشل الناجم عن الصدفة فى سؤال واحد يؤدى الى فرق أكبر فى الدرجة الكلية فى هذه المستويات عنه فى المستويات الدنيا ٠ وحيث أن المفحوصين من ذوى نسب الذكاء العالية (من أى عمر زمنى) عادة مايجيبون على الاسئلة التى تقع فى المستويات العليا من المقياس فان نسب ذكائهم نتيجة لذلك تتضمن مقدارا أكبر من أخطاء القياس وبالتالى تنخفض معاملات الثبات ٠

وتتأكد هذه الحقيقة من ملاحظة جدول انتشار درجات الصورتين م ، ل من طبعة ١٩٣٧ الذى تأكد أنه شكل المروحة فى أى عمر وهذا يعنى أن الصورتين ترتبطان عاليا فى المستويات الدنيا من نســـب الذكاء ، أما فى المستويات العليا فان العلاقة بين الصورتين تتناقص بوضوح ٠

وفى مثل هذه الاحوال يكون من المفضل حساب معامل ارتباط كلى (كمؤشر للثبات) ، ولهذا حسبت معاملات ثبات مختلفة للفئات المختلفة من نسب الذكاء فى كل مرحلة عمرية ٠

وبوجه عام يمكن القول أن مقياس ستانفرد ـ بينيه على درجة عالية

من الثبات فى مختلف المستويات العمرية وفى مختلف نسب الذكاء ، بالرغم من تجانس مجموعات عينة التقنين من حيث العمر الزمنى ، مما يؤثر بالطبع فى معاملات الارتباط عموما ، ويؤدى الى نقصانها ٠٠ وقد تراوحت هذه المعاملات بالنسبة للاعمار ٢½ - ٥½ بين ٨٣ لذوى نسب الذكاء العالية ، ٩١ر٠ لذوى نسب الذكاء المنخفضة ٠ وبالنسبة للاعمار من ٦ - ١٣ تتراوح هذه المعاملات بين ٩١ر، ٩٧ر٠ ولنفس المستويات من نسب الذكاء، وبالنسبة للاعمار من ١٤ - ١٨ كانت معامـــلات الثبات المطابقة للمســـتويين هى ٩٥ر٠ ، ٩٨ر٠

صدق المقياس :

(أ) الصدق المرتبط بالمحكات :تتوافر بيانات كافية عن الصدق المرتبط بالمحكات سواء كان صدقا تلازميا أو صدقا تنبؤيا فى ضوء محك التحصيل المدرسى ٠ فمنذ ظهر مقياس ستانفرد - بينيه عام ١٩١٦ لم يتوقف الباحثون عن حساب معاملات الارتباطات بينه وبين المدرجات المدرسية وتقـــديرات المعلمين ودرجات الاختبارات التحصيلية المقننة ٠ ويتراوح معظم هـــذه المعاملات بين ٤٠ر٠ ، ٧٥ر٠ كما ارتبط بالتفـــوق والتخلف الدارسيين بوجه عام ٠

وقد وجد أن هذا المقياس - كغيره من مقاييس الذكاء - يرتبط ارتباطا عاليا بالأداء فى جميع المواد الدراسية الأكاديمية تقريبا ، الا أن ارتباطه أعلى ما يكون بالمقررات اللغوية (كمواد اللغات او المواد الاجتماعية) ٠ وقد تأكدت هذه النتائج عند اختبار التلاميذ والطلاب فى مختلف المراحـــل التعليمية ، وبالطبع فان معاملات الارتباط تميل الى التناقص مع ارتفاع السلم التعليمى بسبب عوامل الانتقاء وعدم ملاءمة سقف الاختبار مما يؤدى الى نقصان التباين بين المفحوصين ٠

وفى تفسير نسبة الذكاء يجب ان نؤكد ان مقياس ســـتانفرد - بينيه (كغيره من اختبارات الذكاء العام) هو فى الأغلب مقياس للاستعداد المدرسى او الأكاديمى العام ٠ ومشبع الى حد كبير بالمحتوى اللفظى ، وبخاصة فى اختبارات المستويات العمرية العليا ٠ ولذلك فان الأفراد من ذوى النقائص

اللغوية ، او أولئك الذين يتفوقون فى بعض القدرات غير اللغوية (كالقدرات الموسيقية) يحصلون على درجات منخفضة نسبيا فى المقياس ٠ وكذلك فانه توجد بعض المجالات لا يلعب فيها الاستعداد المدرسى او الفهم اللغوى دورا هاما ، وفيها لا تصلح اختبارات الذكاء من نمط ستانفرد - بينيه ٠ ومعنى ذلك أن استخدام الاختبارات بصفة عامة يجب أن يكون فى المواقف التى تلائمها ، والا فان الباحث يتوقع من المقياس مالا يعطيه ٠

(ب) **صدق التكوين الفرضي** : من أهم المحكات التى استخدمت فى انتقاء اختبارات مقياس ستانفرد - بينيه محك تمايز العمر ٠ ومعنى ذلك أن المقياس يقيس قدرات متزايدة مع العمر منذ الطفولة حتى الرشد (فى ظل اطار ثقافى وحضارى معين بالطبع) ٠

وقد استخدم أيضا محك الاتساق الداخلى فى انتقاء الأسئلة ٠ وقد اكدت النتئج وجود نوع من التجانس الوظيفى فى مقياس ستانفرد - بينيه ، بالرغم من التنوع الظاهرى فى المحتوى ، والدليل على هذا أن متوسط معاملات الارتباط بين السؤال والمقياس ككل هو ٦٦ر فى طبعة عام ١٩٦٠ ٠ الا أن من الواضح غلبة الطابع اللفظى على المقياس ، ويتضح هذا من ارتفاع معامل الارتباط بين الأسئلة اللفظية والأداء الكلى فى المقياس عنه فى حالة الأسئلة غير اللفظية ٠

ومن البراهين الاضافية على صدق التكوين الفرضى نتائج الدراسات العاملية العديدة التى أجريت على اختبارات مقياس ستانفرد - بينيه ٠ ويعتمد هذا البرهان على انه لو كانت نسب الذكاء قابلة للمقارنة فى مختلف الأعمار ، فان المقياس يجب أن يكون له نفس التكوين العاملى فى جميع مستويات الأعمار ، بل يجب أن يكون المقياس مشبعا بعامل عام واحد حتى يمكن تفسير نسبة الذكاء تفسيرا واضحا ٠ أما اذا كان المقياس مشبعا بعاملين طائفيين أو اكثر فان نسبة الذكاء الواحدة التى يحصل عليها أفراد متعددون (نسبة ذكاء ١٢٠ مثلا) قد لا تدل على معنى سيكولوجى واحد ٠

وقد تحقق هذا الغرض من نتائج التحليلات العاملية التى اجـــراها ماكنمر (٦٩) لاختبارات مقياس ستانفرد ـ بينيه فى ١٤ مستوى عمريا ابتداء من سن سنتين وحتى سن ١٨ سنة • وكان عدد المفحوصين فى كل تحليـــل عاملى يتراوح بين ٩٩ ، ٢٠٠ ، كما تراوح عدد الاســـئلة التى خضعت للتحليل بين ١٩ ، ٣٥ • وفى كل الأحوال حسبت معاملات الارتباط الرباعية بين المفردات ، وخضعت مصفوفات الارتباط للتحليل العاملى • وتؤكد نتائج ماكنمر أن الأداء فى مقياس ستانفرد ـ بينيه يمكن تفسيره فى ضوء عامل عام واحد ، كما وجدت عوامل طائفية اضافية فى بعض المستويات العمرية ، الا أن اسهام هذه العوامل فى التباين كان قليلا • كما ثبت أن العامل العام المشترك الذى وجد فى المستويات العمرية المتقاربة كان متشابها الى حـــد كبير • ومن أهم النتائج التى توصل اليها ماكنمر أن العامل العام يصبح اكثر لفظية كلما تقدمت الأعمار • ومن ذلك مثلا أن تشبع العامل العام باختبار المفردات ارتفع من ٥٩ر٠ فى سن ٦ سنوات الى ٩١ر٠ فى سن ١٨ سنة •

وأجرى جونز (٦٥ ، ٦٦) دراســــتين عامليتين لأسئلة مقيـــاس ستانفرد ـ بينيه على ٤ مجموعات من الأطفال أعمارهم ٧ ، ٩ ، ١١ ، ١٣ سنة على التوالى ، وكانت كل مجموعة تتكون من ١٠٠ ذكر ، ١٠٠ انثى ، ووجد أنه فى كل عمر يوجد عدد من « القدرات » المتمايزة ولكنها مرتبطة ، ومنها العوامل اللفظية والتذكرية الاستدلالية والمكانية والادراكية • وقـــد اختلفت هذه العوامل وأوزانها النسبية من مستوى عمرى لآخر • وتتفق هذه النتيجة مع نتائج دراسات عاملية أخرى على هذا المقيـــاس يمكن الرجوع اليها ملخصة فى (١٨) •

والواقع أن الأساليب الاحصائية التى استخدمها ماكنمر وجونز من النوع الذى يؤكد دور العامل العام فى الحالة الأولى (دراسات ماكنمر) ، ودور العوامل الطائفية المختلفة فى الحالة الثانية (دراسات جونز) • وحين نتأمل النتائج التى توصل اليها الباحثان نجد أنها تؤكد وجود قدر مشـترك فى القياس ككل ، وهى خاصية من خصـائص بناء المقياس حسب طريقة انتقاء أسئلته التى كانت تؤكد ارتفاع معاملات الارتباط بين السؤال والدرجة الكلية فى المقياس وبالاضافة الى ذلك فان الأداء يتأثر كذلك بعـــدد من القدرات المتخصصة التى يختلف تكوينها تبعا لمستوى العمر الذى نختبره •

اما عن الدراسات التى أجريت فى مصر لتحديد المعالم السيكومترية لمقياس ستانفرد - بينيه فانها قليلة ۰ وتتلخص فى دراسة قام بها لويس كامل مليكه عام ١٩٦٠ (٣٤) للتحقيق من صدق الترتيب المبدئى لاختبارات المفردات فى المقياس ، وذلك بتطبيقها على مجموعات كبيرة من الأفراد ومن مختلف فئات السن (بين ٣ سنوات ، ٣٠ سنة) بلغ عددها ١٦٤٢ مفحوصا ، وفى عام ١٩٦٣ قام محمود الهوارى (٢٦) بدراسة عن ثبات المقياس وصدقه ، استخدم فى حساب الثبات طريقة اعادة الاختبار على ٩٥ تلميذا بالسنة الثالثة الاعدادية (متوسط أعمارهم ١٤ عاما) ، بفاصل زمنى بين شهر وشهرين ، وبلغ معامل الارتباط ٨٣ر٠ اما عينة الصدق فتألفت من ١٦٠ تلميذا بالسنة الثالثة الاعدادية ايضا ، وقد حسب لهم معاملات الارتباط بين درجات المقياس (نسبة الذكاء) وحالات الرسوب والنجاح فى امتحان الشهادة الاعدادية من ناحية ، والمجموع الكلى لدرجات هذا الامتحان من ناحية ثانية ، ودرجات مواد هذا الامتحان من ناحية ثالثة ۰ وقد وجد هذا الباحث أن معامل الارتباط الثنائى بين نسب الذكاء وحالات النجاح والرسوب هو ٥٤ر ، ومعامل الارتباط بين نسب الذكاء والمجموع الكلى لامتحان الشهادة الاعدادية ٨١٨ر٠ اما معاملات ارتباط المقياس بدرجات المواد الدراسية فكانت فى المتوسط ٥٠ر ولم يظهر التحصيل اللغوى (فى اللغة العربية) معاملات ارتباط أعلى من التحصيل اللغوى (فى الرياضيات) كما هو الحال فى الدراسات السابقة ومنها بحث بوند (٥١) الشهير ۰

واعد الباحث بعد ذلك مصفوفة معاملات ارتباط درجات الذكاء ودرجات المواد الدراسية المختلفة وأخضعها للتحليل العاملى فوجد أن تشبع مقياس ستانفرد - بينيه بالعامل العام ٧٨٤ر وهو أعلى التشبعات بهذا العامل ، الا أن هذا ليس برهانا على صدق التكوين الفرض للمقياس لأن باقى المتغيرات كانت فى الواقع درجات للتحصيل المدرسى ، ولم تتضمن المصفوفة أى مقياس آخر للذكاء ۰

وقد ظهرت على نسق الصورة المصرية للمقياس (طبعة ١٩٣٧) صورتان عربيتان أخريان احداهما أردنية أعدها عبد الله زيد الكيلانى (١٢٣) والاخرى سعودية أعدتها آمال أحمد مختار صادق (٩) ، كما ظهرت صورة كويتية لطبعة ١٩٦٠ أعدها رجاء محمود علام وكمال ابراهيم مرسى (١١٦)

كما قام عبد الله عبد العزيز المغيلى (٣٢) بتقنين طبعة ١٩٦٠ على عينة سعودية من ١٧٥ تلميذا بالمرحلة الابتدائية بالرياض تمتد اعمارهم بين ٦ ، ١٢ سنة ، وحسب له الصدق بمحك اختبار رسم الرجل والذى سبق تقنينه على البيئة السعودية وتراوحت معاملات الارتبـاط بين ٥٧ر ، ٨٤ر٠ للاعمار المختلفة ، أما الثبات فحسب بطريقة اعادة الاختبار - بفاصل زمنى طوله أسبوعان ، وتراوح معامل الارتباط بين ٨٨ر ، ٩١ر٠ ، وحسبت له معايير فى صورة متوسطات عمرية تدل على العمر العقلى ، والسؤال الذى لم نجد له اجابة فى هذا البحث هو : كيف استخدم الباحث متوسـطات الأعمار للدلالة على العمر العقلى فى مقياس ستانفرد - بينيه ، ومن المعلوم أن حساب العمر العقلى فى هذا المقياس يقوم على أسس عرضناها بالتفصيل فيما سبق لا تعتمد على فكرة حساب المتوسط والانحراف المعيارى ، ومن الطريف أن الباحث استخدم ضمن معاييره نسـبة الذكاء بالمعنى التقليدى ، (٢٣ : ١٢٩) فى الوقت الذى انقرض فيه هذا المعيار تماما من التراث السيكومترى المعاصر .

وتكاد تكون الدراسة العربية الوحيدة - فى حدود علمنا - التى قامت بتقنين طبعة ١٩٧٢ من هذا المقياس هى دراسة جمال محمد على (١٣) ، وبدأ بالصورة المصرية القائمة والتى أعدها محمد عبد السلام أحمد ولويس كامل مليكه فطبقها على عينة من تلاميذ مرحلة التعليم الأساسى فى مصر (ن = ٢٠٧) تتراوح أعمارهم بين ٦، ١٥ سنة، ثم طبق على الأسئلة الطرق الاحصائية لتحليل الاحتمالات الاختيارية لكل سؤال ، وتحليل قائمة الكلمات وترتب على ذلك تغيير وتعديل وحذف واضافة بعض المفردات لكى تناسـب طبيعة الثقافة والبيئة المصرية . واستعان فى كل ذلك بطبعة ١٩٧٢ من المقياس ، وانتهى به الأمر الى اعداد صورة مقترحة للمقياس طبقها على عينة أخرى من نفس المرحلة (ن = ١١٦) للتأكد من صلاحيتها للاستخدام وتلا ذلك تطبيقها على عينة ثالثة (ن = ١٠٠) لحساب الخصائص السيكومترية لها، فحسب الثبات بطريقة كيودر - ريتشاردسون فبلغ معامل الثبات ٧٤ر ، كما قدر الصدق بحساب معاملات الارتباط بين كل من الدرجات الخـام ونسـب الذكاء الانحرافية محكين هما تقديرات المعلمين ، ومواد التحصيل الدراسى والمجموع الكلى لها فتراوحت معاملات الارتباط بين ٢٠ر ، ٥٥ر وهى معاملات دالة احصائيا بالرغم من انخفاضها الظاهر . وقد قدر الصدق ايضا بطريقة

المجموعات المتضادة ، حيث قورن الأداء الاختبارى لعينة من المتخلفين عقليا والعاديين المساوين لهم فى العمر الزمنى ، ووجدت فروق دالة بينهما لصالح العاديين • وحسب الباحث للصورة المقترحة معاييرها فى صورة نسب ذكاء انحرافية •

اما طبعة ١٩٨٥ من المقياس فلا تكاد توجد دراسة عربية واحـــدة اجريت عليها •

مقاييس وكسلر للذكاء

ظهرت الطبعة الأولى من مقاييس وكسلر عام ١٩٣٩ وعرفت حينئذ باسم مقياس وكسلر ـ بليفو للذكاء Wechsler - Bellevue • وكان أحـــد الأغراض الهامة من اعداد هذه المقاييس توفير اختبار للذكاء يصلح للراشدين وقد أشار وكسلر فى ذلك الحين الى أن الاختبارات التى كانت متاحة حتى ذلك الوقت (ومنها بالطبع مقياس ستانفرد ـ بينيه) معـــدة فى جوهرها للاطفال وتلاميذ المدارس ، وكانت تستخدم مع المراهقين والراشدين عن طريق اضافة بعض الأسئلة الأكثر صعوبة من نفس النوع • ولذلك فان من المشكلات الهامة فى هذه الاختبارات أن محتواها لا يثير اهتمام الراشدين (أى يعوزها الصدق الظاهرى) • واذا لم يتوافر فى المقياس حدا أدنى من الصـــدق الظاهرى يصعب تكوين علاقة طيبة بين الفاحص والمفحوص الراشد •

وبالاضافة الى ذلك فان المبالغة فى تأكيد عامل السرعة فى معظم الاختبارات الفردية يؤدى الى تعطيل أداء الراشدين • وكذلك فان الاهتمام بالتناول الروتينى للنواحى اللفظية تحظى بالاهتمام الأكبر فى هذه المقاييس• ثم انه يستحيل استخدام معيار العمر العقلى مع الراشدين كمـــا أشرنا • واخيرا فان عينات تقنين الاختبارات الفردية (مثل مقياس ستانفرد ـ بينيه) لم تتضمن الا قليلا من الراشدين •

ولمواجهة هذه المشكلات أعد وكسلر مقاييسه للذكاء ، ومن أشـــهرها مقياسه لذكاء الراشدين • والذى يسمى فى الوقت الحاضر Wechsler Adult

أو Intelligence Scale WAIS ، والذى ظهر فى عام ١٩٥٥ ليحل محل مقياس وكسلر ـ بلفيو القديم · وقد ظهرت من المقياس طبعة جديدة معدلة فى أواخر السبعينات باسم WAIS-R لقياس ذكاء الراشدين من سن ١٦ الى سن ٧٤ سنة · وبالطبع يتشابه المقياس الجديد فى الشكل والمحتوى مع المقياس السابق · الا انه تغلب على بعض الصعوبات الفنية الخاصة بمدى تمثيل عينة التقنين · وثبات الاختبارات الفرعية التى يتكون منها المقياس ، ثم ظهرت منه عدة طبعات آخرها ما صدر عام ١٩٨١ باستخدام مقياس وكسلر لذكاء الراشدين المعدل WAIS-R .

وقد ظهر مقياس وكسلر ـ بلفيو الأصلى باللغة العربية فى عام ١٩٥٦ وأعده لويس كامل مليكه ومحمد عماد الدين اسماعيل (٢٨) ومنذ ذلك الحين لم تطرأ عليه تعديلات تذكر ·

وفى عام ١٩٤٦ نشر وكسلر صورة ثانية لمقياس بلفيو الأصلى ، قننت على ألف راشد من الذكور أعمارهم تمتد من ١٨ ـ ٤٠ سنة ، الا ان هذه الصورة سرعان ما فقدت هويتها لأنها صارت أساسا لبناء مقياس آخر لذكاء الأطفال الذى شاع باسم مقياس وكسلر لذكاء الاطفال (WISC) والذى ظهرت منه طبعة معدلة عام ١٩٧٤ باسم (WISC-R) لقياس ذكاء الأطفال من سن ٦ الى سن ١٦ سنة ·

وقدنقل المقياس الأصلى الى اللغة العربية محمد عماد الدين اسماعيل ولويس كامل مليكه عام ١٩٦١ ولا تتوافر عنه بيانات سيكومترية فى اللغة العربية حتى الآن ·

وسرعان ما ظهرت الحاجة الى قياس ذكاء الاطفال فى سن ما قبل السادسة فأصدر وكسلر فى عام ١٩٦٧ اختباره الثالث والهام والذى عرف باسم مقياس وكسلر لذكاء أطفال ماقبل المدرسة والذى عرفباسم (WPPSI) ويقيس ذكاء الأطفال بين ٤ ، ٦½ سنة ·

وتتشابه المقاييس الثلاثة فى البنية العامة حيث يتضمن المقياس مجموعة من الاختبارات الفرعية تشتق منها درجة لفظية Verbal

ومجموعة أخرى تشتق منها درجة أدائية performance ، وترتبط الدرجتان معا فى درجة كلية Full Scale ٠ وتتشابه الاختبارات الفـــرعية فى المستويات العمرية المختلفة الا انها لا تتطابق فهى تختلف فى المحتوى والطول ومستويات الصعوبة والوزن النسبى للسؤال ٠ ووجه التطابق الوحيد هو المعيار المستخدم ، ففى الاختبارات الفرعية تحول الدرجات الخام الى درجات معيارية معدلة متوسطها ١٠ وانحرافها المعيارى ٣ ، اما فى الاختبار الكلى فان الدرجة المعيارية المعدلة موسطها ١٠٠ وانحرافها المعيارى ١٥ ٠

ويوضح الجدول رقم (١٢) الاختبارات الفرعية التى تتالف منهـا المقاييس الثلاثة ويمكن الاستفادة به فى المقارنة ٠

جدول رقم (١٢) الاختبارات الفرعية فى مقاييس وكسلر الثلاثة

WAIS-R مقياس الراشدين	WISC-R مقياس الأطفال	WPPSI مقياس ما قبل المدرسة
	المقياس اللفظى	
المعلومات العامة	المعلومات العامة	المعلومات العامة
الفهم العام	الفهم العام	الفهم العام
الاستدلال الحسابى	الاستدلال الحسابى	الاستدلال الحسابى
المتشابهات	المتشابهات	المتشابهات
المفـردات	المفردات	المفردات
مدى الأرقام	(مدى الأرقام)	(الجمل)
	المقياس الأدائى	
رسوم المكعبات	رسوم المكعبات	رسوم المكعبات
تكميل الصور	تكميل الصور	تكميل الصور
ترتيب الصور	ترتيب الصور	——
تجميع الأشياء	تجميع الأشياء	——
رموز الأرقام	الترميز	بيت الحيوان
——	(المتاهات)	المتاهات
——	——	الرسوم الهندسية

(١) الاختبارات بين القوسين هى اختبارات احتاطية ٠

ونعرض فيما يلي وصفا مفصلا لكل مقياس من المقاييس الثلاثة ، ونبدأ بمقياس الراشدين ثم مقياس الأطفال ، أما مقياس أطفال مرحلة ما قبل المدرسة فسوف نتناوله مع الفئة الخاصة بمقاييس ذكاء الرضع وأطفال ما قبل المدرسة في قسم لاحق من هذا الفصل ٠ ٠

وصف مقياس وكسلر لذكاء الراشدين :

يتكون مقياس WAIS-R من ١١ اختبارا فرعيا ، ٦ منها تصنف في مقياس لفظي verbal scale ، ٥ منها تصنف في مقياس ادائي أو عملي performance ٠ وفيما يلي وصف موجز لهذه الاختبارات الفرعية مع بيان عدد الأسئلة التي يتكون منها كل اختبار في الطبعتين العربية والأجنبية٠

المقياس اللفظي : يتكون المقياس اللفظي من ٦ اختبارات هي :

١ ـ المعلومات العامة : Information ، ويتكون من ٢٩ سؤالا في الطبعة الأجنبية (٢٦ سؤالا في الطبعة العربية منها سؤال تمهيدي) تشمل مدى واسعا من المعلومات العامة التي يفترض شيوعها في ثقافة الراشدين٠ وقد روعي تحرر الاختبار من المعارف التخصصية والأكاديمية ٠

٢ ـ الفهم العام : comprehension ، ويتكون من ١٤ سؤالا في الطبعة الأجنبية (١٠ أسئلة في الطبعة العربية) ، وفي كل سؤال يشرح للمفحوص ما يجب عمله في ظروف معينة ٠ ويتطلب هذا الاختبار قدرة على الحكم العملي والفهم العام Common Sense ٠ وهو يشبه في هذا أسئلة الفهم في مقياس ستانفرد ـ بينيه ، الا أن محتواه أقرب الى اهتمامات الراشدين ونشاطهم ٠

٣ ـ الاستدلال الحسابي arithmetic ويتكون من ١٤ مسألة حسابية في الطبعة الأجنبية (١٠ مسائل في الطبعة العربية) تتفق مع مستوى مادة الحساب في المرحلة الابتدائية ٠ وتعرض كل مسألة شفويا ، ويحلها المفحوص شفويا كذلك دون استخدام الورقة والقلم ، ولكل مسألة زمن محدد ٠

. المتشابهات :Similarities ، ويتكون من ١٢ سؤالا فى الطبعـــة الأجنبية (١٣ سؤالا فى الطبعة العربية) تطلب من المفحوص أن يحدد كيف يتشابه شيئان (البرتقالة والموزة مثلا) ٠

٥ ـ مدى الأرقام (أو اعادة الأرقام) : digit span ، وفيه تعرض على المفحوص قوائم تتكون من ٣ أرقام الى ٩ أرقام ، ويكون عليه اعادتها شفويا ، وفى القسم الثانى من الاختبار يكون على المفحوص اعادة الأرقام بالعكس ٠

٦ ـ المفردات : vocabulary ، وفيه تعرض على المفحوص شفويا وبصريا فى وقت واحد ٤٢ كلمة متزايدة الصعوبة ، ويطلب منه تحديد معنى كل كلمة ٠

المقياس الأدائى أو العملى : يتكون المقياس العملى من ٥ اختبارات هى

٧ ـ رموز الأرقام digit symbol وهو عبارة عن صورة من اختبار التعويض الشفرى code substitution المألوف فى مقاييس الذكـــاء غير اللغوية ٠ ويتكون مفتاح الشفرة من ٩ رموز تتزاوج مع الأرقام التسعة ٠ ويعطى هذا المفتاح للمفحوص ويكون عليه أن يكتب اكبر عدد من الرموز التى تدل على الأعداد الموجودة فى ورقة الاجابة فى زمن لا يتجاوز ٥ر١ دقيقة ٠

٨ ـ تكميل الصــور : Picture completion ويتكون من ٢١ بطاقة فى الطبعة الأجنبية (١٥ بطاقة فى الطبعة العربية) كل منها يحتوى على صورة ناقصة ، ويكون على المفحوص أن يذكر الجزء الناقص من كل صورة ٠

٩ ـ رسوم المكعبات : block design وتشبه مواد هذا الاختبــار ما يوجد فى اختبار Kohs Block Design Test ويتكون من صـــندوق به ١٦ مكعبا ٠ و ٩ بطاقات على كل منها رسم مختلف (اثنتان من هذه البطاقات للتدريب) ٠ وقد لون أوجه المكعبات بالألوان الآتية فى الاصـــل الأجنبى : الأحمر فقط والأبيض فقط والأبيض والأحمر معا ، اما فى الطبعة العربية فهى ملونة بالأزرق فقط والابيض فقط والاصفر فقط والابيض والاحمر معا والازرق

والاصفر معا ٠ويطلب من المفحوص انتاج رسوم على النحو الذى تحدده البطاقات ذات مستويات متزايدة من التعقد تتطلب ما بين ٤ مكعبات ، ٩ مكعبات ٠ولكل بطاقة زمن محدد ٠

١٠ ـ ترتيب الصور : Picture arrangement ، ويتكون كـل سؤال فيه من مجموعة من البطاقات ، عليها صور ، ويطلب من المفحوص ترتيبها بالتتابع الصحيح بحيث تعطى قصة مفهومة ٠ ولكل سـؤال وقت محدد (عدد الأسئلة فى الطبعة الأجنبية ٨ وفى الطبعة العربية ٦) ٠

١١ ـ تجميع الأشياء Object assembly ، ويتبع هذا الاختبار نفس النمط المستخدم فى اختبار Pinter-Paterson الآدائى ٠ ويتكون فى الطبعة العربية من نماذج من الخشب لثلاثة اشياء هى الصبى او المانيكان، والوجه أو البروفيل ، واليد وقد قطع كل منها الى قطع مختلفة ، ويطلب من المفحوص فى كل منها جمع القطع بحيث تكون الشكل الكامل ٠

تصحيح المقياس :

تصحح اختبارات الاستدلال الحسابى ورسوم المكعبات وترتيب الصور وتجميع الاشياء تبعا لسرعة الأداء ودقته على النحو الآتى :

١ ـ الاستدلال الحسابى : تعطى درجة واحدة لكل مسالة تحل حـلا صحيحا فى الوقت المحدد لها ، وتعطى درجة واحدة اضـافية للزمن فى المسالتين ٩ ، ١٠ اذا حلت كل منهما فى حدود ٤٠ ثانية (الزمن العادى لكل منهما ١٢٠ ثانية) ، وتعطى درجتان اضافيتان اذا حلت المسالة فى حدود ١٥ ثانية ، يحسب الزمن بعد انتهاء المفحوص من قراءة المسـالة (النهاية العظمى ١٤) ٠

٢ ـ رسوم المكعبات : تقدر الرسوم حسب الـدقة والزمن الاقصى ، وتعطى درجات اضافية فى حالة اتمام الرسم فى زمن أقل ٠ وقد أعد جدول لاعطاء تقديرات حسب زمن الانجاز (النهاية العظمى ٤٢) ٠

٣ ـ ترتيب الصور : تقدر الدرجة فى هذا الاختبار على اساس دقة الترتيب والزمن المستغرق فيه ، واعدت جداول لهذا الغرض بحيث تصحح كل من المجموعات الثلاث الأولى على أساس الكل او لا شيء (صواب أو خطأ) ٠ اما المجموعات الثلاث الأخيرة فتعطى تقـديرات جزئية على كل ترتيب لا يطابق الترتيب الصحيح ، ولكن يمكن قبوله ٠ وبالنسبة للمجموعتين الأخيرتين تعطى تقديرات اضافية على الترتيب الصحيح اذ تم فى زمن معين، وتحسب الدرجـة النهائية فى هذا الاختبار بجمع التقـديرات الجزئية فى المجموعات المختلفة (والنهاية العظمى ٢١) ٠

٤ ـ تجميع الأشياء : يصحح الشكل الأول (الصبى) حسـب الدقة فقط ، أما الشكلان الآخران (الوجه واليد) فيصحح كل منهما حسب الدقة والزمن معا ٠ وفى الحالة الاخيرة تكون الدرجة فى مواد الاختبـار هى المجموع الكلى لتقديرات الدقة والزمن جميعا ٠ وتوجد بكراسة التعليمات بروتوكولات خاصة لتصحيح كل شكل حسب الدقة ، وجدول للتصحيح على أساس الزمن (النهاية العظمى ٢٦) ٠

أما اختبارات المعلومات العامة ، ومدى الأرقام (أو اعادة الأرقام)، وتكميل الصور ، ورموز الأرقام فان الدرجة الكلية فيها هى مجموع الاجابات الصحيحة حسب مفتاح التصحيح المبين فى كراسة التعليمات ، والنهايات العظمى لهذه الاختبارات الأربع على التوالى هى : ٢٥ ، ١٧ ، ١٥ ، ٦٧ ٠

وقد قام لويس كامل مليكه (٣٥) باعداد نماذج تصـحيح خاصـة باختبارات الفهم العام ، والمتشابهات ، والمفردات ، نشرها فى كراسة منفصلة عام ١٩٦٠ ، وفى هذه النماذج تعطى لأسئلة اختبارى الفهم العام والمتشابهات درجتان أو درجة واحدة أو صفر حسب درجة التعميم فى الاجابة ونوعه ، وبهذا تكون النهاية العظمى لكل من هذين الاختبارين هى : ٢٠ ، ٢٤ على التوالى ٠ أما اختبار المفردات فان أسئلته تعطى فى نماذج التصحيح درجة واحدة أو نصف درجة أو صفر (النهاية العظمى ٢٤) ٠

المعـــايير :

طبعاته المختلفة من العينات القلائل التى يتوفر فيها شرط التمثيل ، فقـــد كانت تتالف من ١٧٠٠ شخصا فى المقياس الأصلى، اما المقياس المعدل الذى صدر عام ١٩٨١ فقد اعتمد على عينة تقنين قوامها ١٨٨٠ تتساوى فيها اعداد الرجال والنساء موزعين فى مستويات عمرية سبع بين ١٦ سنة و ٧٤ سنة اختيروا بحيث تتوافر فيهم خصائص التعداد العام للســكان فى الولايات المتحدة الذى أجرى عام ١٩٧٠ من حيث المنطقة الجغرافية والبيئة الريفية والحضرية والأصل العنصرى (وخاصة البيض والزنوج) ، والمستوى المهنى والمستوى التعليمى ، واقتصرت العينة على « الراشــدين الأســوياء » ، واستبعدت حالات التخلف العقلى ، والمصابون باصابات المخ ، والذين يعانون من اضطرابات سلوكية وانفعالية أو نقائص جسمية ، كما استبعدت الحالات التى لا تستطيع استخدام اللغة الانجليزية ٠

وقد حولت الدرجات الخام فىالاختبارات الفرعية الىدرجات معيارية معدلة متوسطها ١٠ وانحرافها المعيارى ٣ ٠ وكانت هذه الدرجاتالموزونة Scaled Scors مشتقة من جمــاعة مرجعية عددها ٥٠٠ حالة شــملت جميع المفحوصين من سن ٢٠ سنة ، ٣٤ سنة فى عينة التقنين ٠ وهكذا أصبحت الدرجات الفرعية فى صورة وحدات قابلة للمقــارنة ، ثم حسبت درجات المقياس اللفظى ، والمقياس الأدائى ، والمقياس الكلى بجمع الدرجات الموزونة فى الاختبارات اللفظية الست ، والاختبارات الأدائية الخمس ، والاختبارات الأحد عشر جميعا على التوالى ٠ ثم حولت هذه الدرجات الثلاثة الى نسبة ذكاء انحرافية متوسطها ١٠٠ وانحرافها المعيــارى ١٥ فى كل مجموعة عمرية ، وهكذا يتحدد الوضع النسبى للمفحوص بمقارنته بالأشخاص من نفس مستواه العمرى ٠

وقد استخدمت نفس الاجراءات فى تقنين الطبعة العربية حيثاستخدمت فى اعداد الدرجات الموزونة ثلاث فئات عمرية هى :

١ ـ من ٢٠ الى أقل من ٢٥ سنة وعدد أفرادها ١١٤ ٠

٢ ـ من ٢٥ الى أقل من ٣٠ سنة وعدد أفرادها ١١٢ ٠

٣ ـ من ٣٠ الى أقل من ٣٥ سنة وعدد أفرادها ١٠٢ ٠

وقد استخدمت هذه العينة البالغ عددها ٢٢٨ مفحوصا فى اعداد نماذج ومعايير التصحيح لاختبارات الفهم والمتشابهات والمفردات بالاضافة الى اعداد جدول الدرجات الموزونة وجداول لنسب الذكاء الانحرافية لكل فئة من فئات السن (٣٥) ٠

وفى تصحيح هذا القياس تعطى للمفحوص درجات ثلاث : نسبة الذكاء الكلية ونسبة الذكاء اللفظية ونسبة الذكاء العمليـــة ٠ ويعطى فى الطبعة الأجنبية درجة رابعة تدل على ما يسميه وكســـلر « معامل الكفـــاءة » efficiency quotient أو (E.Q.) ، ويتحـــد بتقدير نسـبة ذكـاء الفرد فى الفئة من ٢٠ الى ٢٤ (فى اختبار وكسلر بلفيو طبعة ١٩٣٩) وفى الفئة ٢٥ ـ ٢٩ (فى اختبار وكسلر لذكاء الراشدين طبعة ١٩٥٥) ، بصرف النظر عن العمر الحقيقى للمفحوص ٠ ويعتمد هذا المعامل على أساس أن الفئة من ٢٠ ـ ٢٤ (فى تقنين ١٩٣٩) والفئة من ٢٥ ـ ٢٩ (فى تقنين ١٩٥٥) هما اللتان وصلتا الى أعلى مستويات الأداء فى المقياس ٠ ومعنى ذلك أن معامل الكفاءة طريقة للحكم على اداء الفرد فى ضوء متوسط أداء الأفراد من المستوى العمرى الذى يصل فيه الأداء الاختبارى اقصاه ٠

ويذكر لويس كامل مليكه (٣٥ : ٦١) أنه يمكن الحصول على « معامل الكفاءة بصورة مؤقتة عن طريق استخدام جداول نسب ذكاء للفئة من ٢٠ ـ ٢٤ من العينة المصرية انفة الذكر ، لأن هذه الفئة حققت بالفعل أعلى مستوى من الأداء فى المقياس ٠

تفسير المعايير :

اعد وكسلر عام ١٩٤٤ فئات الذكاء التى يمكن أن تستخدم فى أغراض التصنيف الكلينيكى ٠وكان أساس التصنيف هو حـــدود الخطأ المعيارى للمتوسط ٠ فالشخص الذى تكون درجته اقل من المتوسط بمقدار يبلغ ثلاثة امثال الخطأ المعيارى أو اكثر يعد من فئة ضعاف العقول ٠ أما الذى يحصل على درجة تقع ما بين خطأين معياريين وثلاثة اخطاء معيارية فيعد من قبيل الحالات الهامشية ، وهكذا ٠ ويبين الجدول رقم (١٣) أسس تصـــنيف مستويات الذكاء فى مقياس وكسلر ٠

جدول رقم (١٣) تصنيف مستويات الذكاء فى مقياس وكسلر

التصنيف	حدود الخطأ المعيارى	حدود نسبة الذكاء الانحرافية	النسبة المئوية فى الأصل الاحصائى
ضعاف العقول	من ٣ فأقل	٦٥ فأقل	٢ر٢
الحالات الهامشية	من ـ ٢ حتى ـ ٣	من ٦٦ حتى ٧٩	٦ر٧
الأغبياء	من ـ ١ حتى ـ ٢	من ٨٠ حتى ٩٠	١ر١٦
المتوسطون	من ـ ١ حتى + ١	من ٩١ حتى ١١٠	٠ر٥٠
أعلى من المتوسط	من + ١ حتى + ٢	من ١١١ حتى ١١٩	١ر١٦
الممتازون	من + ٢ حتى + ٣	من ١٢٠ حتى ١٢٧	٦ر٧
الممتازون جدا	من + ٣ فاكثر	من ١٢٨ فاكثر	٢ر٢

ثبات المقياس :

حسبت معاملات الثبات للطبعـــة الأجنبية من مقياس وكســـلر لذكاء الراشدين لمجموعات من فئات الأعمار ١٨ ـ ١٩ ، ٢٤ ـ ٣٤ ، ٤٥ ـ ٥٤ اختيرت من عينة التقنين على أساس تمثيلها لمدى العمر الذى تشمله هذه العينة • واستخدمت طريقة التجزئة النصفية (بعد تصحيح سبيرمان ـ براون) فى حساب ثبات كل اختبار فرعى فيما عدا اختبـارى رموز الأرقام ومدى (اعادة) الأرقام حيث استخدمت معهما طريقة الصور المتكافئة لعدم صلاحية طريقة التجزئة النصفية لهما بسبب غلبة عامل السرعة الشديدة عليهما •

ويلاحظ أن معاملات الثبات للمقياس الكلى قد بلغت فى المجموعات الثلاثة ٩٧ر ، وبلغت فى المقياس اللفظى ٩٦ر ، وفى المقياس الادائى ٩٤ر٠ ومعنى هذا أن المقاييس الثلاثة على درجة عالية من الثبات فى ضوء الاتساق الداخلى •

أما فى الطبعة العربية فقد حسب فرج عبد القادر طه (٣٠ : ٧٨) ثبات المقياس بطريقة اعادة الاختبار وطريقة التجزئة النصفية على عينة من عمال الصناعة ، وفى الطريقة الأولى أعيد تطبيق المقياس على ٤٠ حالة من

العمال فتراوحت معاملات الثبات بين ٦ر ،٩ر٠ بالنسبة لكل من نسب الذكاء والاختبارات الفرعية • اما طريقة التجزئة النصفية فقد حسبت لدرجات ٧٠ عاملا وتراوحت معاملات الثبات الناتجة عنها بين ٥ر ، ٩ر٠ لنسب الذكاء والاختبارات الفرعية • ولم يوضح الباحث كيف استخدم طريقة التجزئة النصفية لاختبارى رموز الأرقام واعادة الأرقام بالرغم من عدم صلاحية هذه الطريقة لهما كما أشارت نتائج بحوث التقنين التى اجـــريت على المقياس (٤٩ : ٢٤٧) •

صدق المقياس :

يفضل وكسلر (٨٣) فى تناوله لصدق مقياسه ان يؤكد « صـــــدق المحتوى » • ومن رأيه أن الاختبارات الأحد عشر المتضمنة فيه تتفق مع تعريف الذكاء ، وأن اختبارات مماثلة استخدمت فى مقاييس الذكاء السابقة، وثبت جداوها فى الأغراض الكلينيكية وهذا القول اقرب الى صدق التكوين الفرضى ولكن بدون شواهد امبريقية • وبالاضافة الى ذلك يذكر وكسلر(٨٣) وماتارازو (٦٨) بعض البيانات عن الصدق التلازمى للمقياس • ومن ذلك المقارنة بين المجموعات التعليمية والمهنية المختلفة ، بالاضافة الى حساب معاملات الارتباط مع تقديرات الأداء على العمل والتقديرات المدرســية ، وكانت أغلب الفروق بين المجموعات فى الاتجاه المتوقع ، ومن ذلك ان درجات العمال أعلى فى الاختبارات العملية ، وان درجات الأفـــراد فى المهن غير اليدوية أعلى فى الاختبارات اللفظية • اما عن معاملات الارتباط بالمحكات المهنية والمدرسية فقد أعطى المقياس اللفظى قيما أعلى من المقياس الكلى أما المقياس الأدائى فقد أعطى ادنى القيم • وفى كل الأحـــوال كانت معاملات ارتباط المقياس اللفظى أدنى من معاملات مقياس ستانفرد ـ بينيه • أما عن الصدق التنبؤى للمقياس فقد ثبت أن المقياس يتنبا بدرجة مقبولة بالضعف العقلى كما تؤكد بحوث جيرتن وزملائه (٦١ ، ٦٢) •

وقد حسبت معاملات الارتباط بين مقياس وكسلر وغيره من الاختبارات والمقاييس المعدة لقياس الذكاء ، وكانت النتائج كما يلى :

١ ـ معامل الارتباط بين المقياس ومقياس ستانفرد ـ بينيه ٨٠ر٠ وهو أعلى فى المقياس اللفظى منه فى المقياس الكلى والمقياس الأدائى •

٢ ـ معاملات الارتباط بين المقياس واختبـــارات الذكاء الجماعية تتراوح بين ٤٠ر ، ٨٠ر٠

٣ ـ معامل الارتباط بين المقياس الأدائى واختبار منيســوتا للوحات الأشكال ٧٢ر٠

٤ ـ معامل الارتباط بين المقياس الأدائى واختبار زافن للمصفوفات المتتابعة ٧٠ر٠

٥ ـ معامل الارتباط بين المقيـــاس الأدائى واختبـــار بينت للفهم الميكانيكى ٣٥ر٠

صدق التكوين الفرضى : أجريت عدة تحليلات عاملية للاختبـارات الفرعية التى يتألف منها مقياس وكسلر ٠ ومن أشهر هذه الدراسات بحوث كوهن (٥٥ ، ٥٦) على مصفوفات ارتباطية لمجموعات عمرية اربعة من عينة التقنين هى : (من ١٨ ـ ١٩ ، ومن ٢٥ ـ ٣٤ ، ومن ٤٥ ـ ٥٤ ، ومن ٦٠ ـ ٧٥ +) ٠ وقد وجد الباحث أن الاختبارات الفرعية جميعا تشترك فى عامل واحد يسهم بحوالى ٥٠٪ من التباين الكلى للبطارية ٠ وبالاضافة الى ذلك استخراج ثلاثة عوامل طائفية هى :

١ ـ عامل الفهم اللفظى وتشبع باختبارات المفردات والمعلومات العامة والفهم العام والمتشابهات ٠

٢ ـ عامل التنظيم الادراكى وتشبع باختبارات رسوم المكعبات وتجميع الأشياء ، وهو عبارة عن مركب من عامل السرعة الادراكية والتصـــور البصرى المكانى اللذين اكدتهما البحوث العاملية الأساسية وبخاصة بحوث ثرستون ٠

٣ ـ عامل الذاكرة وتشبع باختبارات الاستدلال الحسابى ومدى (اعادة) الأرقام ، ويتضمن كلا من الذاكرة الصماء المباشرة للمواد الجديدة واستدعاء المواد التى سبق تعلمها ٠

ومن الطريف أن نذكر أن نتائج هذا البحث لم يدعم الاجـــراء الذى يستخدم فى تصحيح هذا المقياس أى تقسيمه الى مقياس لفظى ومقياس عملى ادائى لكل منهما نسبة ذكاء منفصلة · ورغم ان استخدام نســـب الذكاء الكلية يدعمه وجود عامل عام مشترك فى جميع الاختبارات ، فان عامل الفهم اللفظى لم يظهر الا فى ٤ اختبارات لفظية فقط من ٦ اختبارات ، اما عامل الذاكرة فقد وجد فى الاختبارين اللفظيين الباقيين بالاضافة الى اختبارات فرعية أخرى من المقياسين عند الراشدين الأكبر ســـنا · أما عامل التنظيم الادراكى فكانت تشبعاته الدالة فى اختبارين فقط من الاختبـــارات الأدائية الخمس ، وأما الاختبارات الأدائية الثلاث الباقية فهى فيما يبدو لها عواملها الخاصة التى لا تشترك مع الاختبارات الأخرى المتضمنة فى بطارية المقياس·

وقد قام محمد عماد الدين اسماعيل (٤٢) بتحليل عاملى لمقياس وكسلر (الطبعة العربية) توصل منه الى عامل عام واحد مشترك فى جميع اختبـارات المقياس لفئة السن من ١٦ ــ ١٩ ، وكانت أكثر الاختبارات تشبعا به اختبار المعلومات العامة ثم اختبار الفهم العام ، وتراوحت تشبعات الاختبـــارات الفرعية به بين ٩٥ر ، ٦٤ر· ومن الطـــريف أن الباحث يطابق بين العامل العـــام والـــذى وجـــده فى بطـــارية وكســـلر والعـــامل العـــام عند سبيرمان ، أما بالنسبة لفئـــة السن من ٢٠ ــ ٢٤ فقـــد وجـــد عاملين : عامل عام مشترك فى جميع الاختبـــارات ، وعامل ثنائى القطب تشبع به تشبعات موجبة كاختبارات المعلومات والفهم والمتشابهات والمفردات ، وتشبع به تشبعات سالبة كاختبارات ترتيب الصور وتكميل الصور وتجميع الأشياء ويمكن اعتباره عاملا يميز بين الاختبارات اللفظية فى مقابل الاختبارات غير اللفظية ·

واخيرا نذكر أن البحوث العاملية العديدة التى قام بها سوندرز (٧٤ ، ٧٥ ، ٧٦ ، ٧٧) ، مستخدما مصفوفات ارتباط الأسئلة الفردية للاختبارات الفرعية للمقياس ، توصلت الى ١٠ عوامل على الأقل ، ولم يجـد الباحث تطابقا كاملا بينها وبين الاختبارات الفرعية للمقياس ، بل تأكد أن بنية بعض الاختبارات الفرعية على درجة كبيرة من « التعقد العاملى » ، وأن بعض العوامل تتعدى أكثر من اختبار فرعى واحد ·

ولم يظهر لهذا المقياس صدى يذكر فى البيئات العربية الأخرى بعد

ظهور الصورة المصرية للمقياس ، باستثناء ما قام به فرج عبد القادر طه وصلاح احمد مرحاب (٣١) باعدادهما الصورة المغربية للمقياس ، وقاما بتقنينها على عينة من ١٥٠ مفحوصا من الجنسين ومن الاعمار من ٢٠ ـ ٣٤ سنة ، وحسب الثبات عن طريق اعادة الاختبار للاختبارات الأحد عشر وتراوحت معاملات الثبات بين ٧٧ر ، ٩٦٨ر ، وقدر الصدق باستخدام مجموعتين متضادتين من ضعاف العقول والأذكياء وكانت الفروق بينهما دالة فى المقاييس الثلاثة ، وأعدت جداول الدرجات الموزونة ونسب الذكاء الثلاثة ٠

وصف مقياس وكسلر لذكاء الأطفال :

يتكون مقياس وكسلر لذكاء الأطفال Wechsler Intelligence Scale for Children أو WISC من الأسئلة السهلة التى كان يتألف منها ، مقياس وكسلر ـ بلفيو الأصلى ٠ ويقسم المقياس ـ كما هو الحال فى مقياس الراشدين ـ الى قسمين : أحدهما لفظى والآخر عملى ، والمجموع الكلى للاختبارات الفرعية ١٢ اختبارا منها اختباران احتياطيان أو اضافيان عند الضرورة ، على النحو الآتى :

القسم اللفظى	القسم العملى (الأدائى)
(١) المعلومات العامة	(٦) تكملة الصور
(٢) الفهم العام	(٧) ترتيب الصور
(٣) الاستدلال الحسابى	(٨) تصميم (رسم) المكعبات
(٤) المتشابهات	(٩) تجميع الأشياء
(٥) المفردات	(١٠) الترميز
(اعادة الأرقام)	(المتاهات)

وقد اعتبر وكسلر الاختبارات الاحتياطية هى تلك التى ارتبطت أدنى الارتباطات ببقية المقياس ، وعلى ذلك فان اختبار مدى (اعادة) الأرقام فى المقياس اللفظى واختبار المتاهات فى المقياس العملى من هذا النوع ٠ ويتطابق اختبار الترميز مع اختبار رموز الأرقام فى مقياس الراشدين فيما عدا غلبة الأسئلة السهلة عليه ٠ اما الاختبار الوحيد الذى ليس له نظير فى

مقياس الراشدين فهو اختبار المتاهات ، ويتكون من ٨ متاهات ورقية متزايدة الصعوبة ، وتصحح فى ضوء الزمن وعدد الأخطاء · وفى راى وكسلر أن اختبارى الترميز والمتاهات يمكن أن يحل احدهما محل الآخر ، ويتطلب اختبار الترميز زمنا أقل من اختبار المتاهات ، وقد يفضله بعض الفاحصـــين لهذا السبب ·

تصحيح المقياس وتفسير الدرجات :

ولا يختلف مقياس الأطفال فى اجراءات تصحيحه من مقياس الراشدين الا اختلافات طفيفة · وتتلخص هذه الاجراءات فى تحويل الدرجات الخام فى كل اختبار فرعى الى درجات معيارية معدلة فى اطـــــار المجموعات العمرية التى ينتمى اليها المفحوص · وقد أعد وكسلر جداول لهذه الدرجات الموزونة لفئات عمرية مقدارها أربعة شهور ما بين سن ٥ سنوات حتى سن ١٥ سنة · وتشبه الدرجات الموزونة هنا درجات مقياس الراشدين فى أن متوسطها ١٠ وانحرافها المعيارى ٣ ، ثم تجمع الدرجات الموزونة وتحول الى نسب ذكاء انحرافية متوسطها ١٠٠ وانحرافها المعيارى ١٥ سواء لما قياس اللفظى أو المقياس الأدائى أو المقياس الكلى · وكذلك اهتم وكسلر فى المقياس الأصلى باقتراح بعض الطرق الاحصائية لحساب الأعمار العقلية المقابلة لدرجات مقياس الأطفال · ورغم ان معيار العمر العقلى ليس من خصائص نسب الذكاء الانحرافية التى تتألف منها معايير المقياس الا انه كان فى ذلك يواجه مطلبا عمليا ملحا نتيجة لشيوع مفهوم العمر العقلى بين خاصة الناس وعامتهم بالنسبة لميدان قياس الذكاء · وقد أعدت بالفعل جداول لتحويل الدرجات الخام فى المقياس الى أعمار عقلية · الا أن هذه الحاجة لم تعد ضرورية فى الوقت الحاضر ولذلك نجد في المقياس المعـــدل (WISC-R) يستبعد هذه الضرورة ، وبالاضافة الى ذلك فان المقياس المعدل يتخفف كثيرا من الأسئلة التى لا تتناسب مع الأطفال وأحل محلها أسئلة مشتقة من خبرات الطفولة ·

كما زيد طول بعض الاختبارات الفرعية لزيادة الثبات ، كما طرأت تحسينات على طرق تطبيق المقياس واجراءات تصحيحه ، وقد أدى ذلك الى اعطاء متوسطات نسب ذكاء تقترب كثيرا من القيم التى نحصل عليهـا من

مقياس ستانفرد ـ بينيه (معايير ١٩٧٢) ، وبصفه عامة لوحظ أن معايير المقياس المعدل اعلى من معايير المقياس الأصلى .

وتتالف عينة تقنين مقياس وكسلر المعدل للاطفـــال (WISC-R) من ١٠٠ طفل و ١٠٠ طفلة فى الأعمار من ٥ر٦ سنوات الى ٥ر١٦ سنة ، وكان المجموع الكلى للعينة ٢٢٠٠ مفحوصا صنفت على اساس الاحصاء العام للسكان فى الولايات المتحدة الذى اجرى عام ١٩٧٠ وفى ضوء نفس المتغيرات الذى ذكرناها بالنسبة لمقياس الراشدين ، وكانت جميعا من الاطفال الأسوياء .

ثبات المقياس : حسبت معاملات الثبات بطريقة التجزئة النصفية لكل اختبار فرعى فى مقياس وكسلر للاطفال ولكل من المقياس اللفظى أو المقياس العملى ، والمقياس الكلى وقد حسب الثبات بالتجزئة النصفية لكل مجموعة عمرية من المجموعات الاحدى عشرة الا ان الثبات بطريقة اعادة الاختبار حسب لثلاث مجموعات فقط هى : (أ) من ½٦ ـ ½٧ (ب) من ½١٠ ـ ½١١ (حـ) من ½١٤ ـ ½١٥ وقد بلغ متوسط معاملات الثبات بالتجزئة النصفية ٩٤ر للمقياس اللفظى ، ٩٠ر للمقياس العملى ٩٦ر للمقياس الكلى بينمـا بلغت هذه المعاملات بطريقة اعادة الاختبار ٩٣ر ، ٩٠ر ٩٥ر على التوالى .

اما معاملات ثبات الاختبارات الفرعية فكانت فى الطبعة المعدلة (طبعة ١٩٧٤) اعلى منها فى الطبعات السابقة ، فتراوحت معاملات الثبات لمختلف الأعمار بالتجزئة النصفية بين ٧٠ر و ٨٦ر ، وبطريقة اعادة الاختبار بين ٦٥ر ، ٨٨ر .

صدق المقياس : لاتتضمن كراسات تعليمات مقياس وكسلر للاطفال بيانات عن صدقه ، ولو أن بعض الشواهد تدل على ذلك فجداول المعايير لكل اختبار فرعى تظهر تمايزا فى العمر دون أن تقدم البيانات للقارىء فى ضوء هذا المحك. وقد قام عدد من الباحثين بحساب الصدق التلازمى للمقياس باستخدام محك التحصيل المدرسى وتراوحت ماملات الارتباط بين ٥٠ر ، ٦٠ر وكما هو متوقع كانت معاملات المقياس اللفظى أعلى. كما حسبت معاملات ارتباط مقياس ستانفرد ـ بينيه فبلغت ٧٣ر بالنسبة للمقياس الكلى ، وكانت معاملات المقياس اللفظى مرة أخرى أعلى .

وتقدم كراسة تعليمات المقياس المعدل معلومات اضافية عن معاملات الارتباط بين الاختبارات الفرعية بعضها وبعض من ناحية، والمقياس اللفظى والمقياس العملى والمقياس الكلى من ناحية أخرى · وقد حسبت هذه المعاملات منفصلة لعينة قوامها ٢٠٠ مفحوص فى كل مجموعه عمرية من المجموعات الاحدى عشرة فى عينة التقنين · وقد بلغ متوسط معاملات ارتباط المقياس اللفظى والعملى ٦٧ر· ومعنى هذا أن جزئى المقياس بينها قدر مناسب من الاشتراك الا أن انخفاض معامل ارتباطهما نسبيا يبرر الابقاء عليهمـــا كمقياسين منفصلين ·

أما بحوث التحليل العاملى التى أجريت على المقياس بهدف تحقيق صدق التكوين الفرضى فقد استخرجت عوامل تشبه ما وجد فى تحليـــل مقياس الراشدين وهى : العامل العام ، والفهم اللفظى · والعامل المكانى الادراكى· وعامل الذاكرة · وتأكد أن عامل الفهم اللفظى يرتبط ارتبـــاطات جوهرية بالأختبارات الفرعية الخمسة فى المقياس اللفظى ، وان العـــامل المكانى الادراكى يرتبط ارتباطا دالا باختبارى رسم المكعبات وتجميع الأشياء وهما من الاختبارات الفرعية فى المقياس الأدائى العملى ·

تقويم المقياس : يمكن القول أن مقياس وكسلر للاطفال مقياس جيـــد لذكاء الأطفال ويتفق فى الخصائص مع كثير من الاختبارات الفردية · وكانت انستازى ترى فى طبعة ١٩٦٨ من كتابها (٤٩) أن الأساس المنطقى لهـــذا المقياس ظاهر التناقض · لقد ظهر مقياس وكسلر بلفيو الأصلى عام ١٩٣٩ لمقابلة الحاجة الى اختبار للراشدين لا يوفرها مقياس ستانفورد ــ بينيه أو أى مقياس آخر يعتمد على مجرد (رفع مستوى) مقاييس الأطفال ، وبعدما حقق وكسلر هذا الهدف سعى الى اعداد مقاييس للاطفال تعتمد فى جوهرها على (خفض مستوى مقاييس الراشدين) بل انه سار بالتناقض الى نهايته حينما أعد مقياسا لأطفال ما قبل المدرسة · الا أن هذا النقد لم يعد له ما يبرره بعد أن ظهرت الطبعة المعدلة من المقياس عام ١٩٧٤ وتضمنت محتوى « طفولى » تماما ــ الا أن أهم ما يلفت النظر فى هذا المقياس انه لم يحظ بدراسات تستحق الذكر عنه باللغة العربية ولا تكاد تتوافر بيانات عن معالمه السيكومترية يمكن الاعتماد عليها ·

المقياس البريطاني للقدرة :

تولى مسئولية اعداد هذا المقياس ثلاثة من علماء النفس البريطانين هم اليـــوت C.D. Elliott وموراى D.J. Murray وبيرســـون L.S. Pearson وظهر باسم المقياس البريطاني للقدرة British Ability Scale وقامت بنشره المؤسسة القومية للبحـــوث التربوية (NFER) .

ويتألف هذه المقياس من ٢٤ مهمة يجد فيها معظم المتخصصين فائدة كبرى في التشخيص لمستوى واسع من الأعمار ابتداء من سن ٢½ سنة وحتى ١٧ ســنة ·

وتصنف المهام تبعا لمدى الأعمار ، فالمهام الخاصة بالمزاوجة الادراكية perceptual matching تقتصر على المدى العمرى من ٢½ الى ٩ سنوات، بينما لا تقاس « السرعة في تجهيز المعلومات » قبل سن ٨ سنوات · ومن أمثلة المهام التي تقيس القدرة الأخيرةأن يعرض على المفحوص خمسة اعداد يتألف كل منها من ثلاثة أرقام ويطلب منه أن يختار اكبرها · وتوجد مهام أخرى لقياس الذاكرة والاستدلال الصورى بمعناه عند جان بياجيه بالاضافة الى الاستدلال الاجتماعى بالاضافة الى المعرفة الأساسية بالقراءة والكتابة والسهولة في التعامل مع الأعداد ·

والغرض الرئيسى من هذا الاختبار هو رسم بروفيل المفحوص ·ويمكن للفاحصين استخدام بعض الاختبارات الفرعية دون البعض الآخر · ويرى مؤلفو الاختبار أن بعض الاختبارات الفرعية يمكن الحصول على متوسط درجاتها وتقدير « نسبة ذكاء عامة » منها والبعض الآخر يعطى « نسبة ذكاء بصرية » والبعض الآخر يعطى « نسبة ذكاء لغوية » وهو في هذا أشـــبه بمقياس وكسلر ، وفي نفس الوقت فيه بعض خصائص مقياس ستانفرد - بينيه ولعل من أهم السمات الهامة في هذا المقياس ان بعض اختباراته الفرعية يمكن أن تعطى وأن مفرداتها يمكن أن تختار في ضوء اداء المفحوص ذاته · ومن ذلك مثلا أن اختبار « تصميم المكعبات » فيه يتألف من ١٦ نمطا ، ويمكن للطفل الصغير أن يعطى الأنماط الأحد عشر الأولى بينما يبدأ المراهق من النمط السادس حتى النمط الأخير · ويقدر مستوى قدرة المفحوص من خلال

سجل تراكمى لنجاحه وفشله بطريقة راش Rasch • ويبدو لنا أن هذا المقياس الجديد لا يزال فى حاجة الى مزيد من البحث ، كما ندعو الباحثين المصريين الى محاولة تطويعه للثقافة المصرية خاصة والعربية عامة•

بطارية كوفمان :

فى عام ١٩٨٣ أصدر كوفمان وزوجته اختبارا هاما باسم بطارية تقويم الأطفال لكوفمان Kaufman Assessment Battery for children بحيث تصلح للاطفال من سن ٢½ الى ١٢ سنة وتتالف البطارية من ١٦ اختبارا ، الا أنها مثل المقياس البريطانى للقدرة ، تخصص بعض الاختبارات الفرعية للاطفال الصغار والبعض الآخر للاطفال الكبار •

وتقوم الأسس النظرية للبطارية على الاتجاهات الحديثة فى مجال سيكولوجية الجهاز العصبى وعلم النفس المعرفى ، ولهذا يمكن الحصول منه على أربع درجات ، كلية ، رئيسية هى :

١ ـ درجة التجهيز التتابعى Sequential Processing ويمكن الحصول عليها من أداء المفحوص لمهام الذاكرة والتى تتطلب منه اعادة انتاج سلسلة من الأعداد بعد الاستماع اليها ، أو اصدار سلسلة من حركات، اليد بعد رؤيتها أو لمس رسوم يسهل التعرف عليها بنفس الترتيب الذى يستخدمه الفاحص فى تسمية الاشياء المصورة فى هذه الرسوم •

٢ ـ درجة التجهيز المتانى Simultaneous Processing ويمكن الحصول عليها من أداء مهام تشبه ما تتضمنه اختبارات مشابهة لترتيب الصور فى مقياس وكسلر ، ولاختبار المصفوفات ، واختبار الاغلاق (الجشطالتى) • ومن الاختبارات الجديدة المستحدثة فى هذه البطارية اختبار يسمى « النافذة السحرية » وفيه يحرك الفاحص صورة بحيث تمر فى نافذة ضيقة ويطلب من المفحوص أن يتعرف على الشيء المصور •

٣ ـ الدرجة المركبة للتجهيز العقلى وتمثل متوسط الدرجتين السابقتين • وتمثل الدرجة التى يحصل عليها المفحوص فى الاختبارات

الفرعية التى يتم تطبيقها عن طريق التمثيل الصامت ما يسمى « الدرجة غير اللفظية » ٠

٤ ـ درجة التحصيل ويمكن الحصول عليها من اداء المفحوص على المهام التى تعتمد مباشرة على الخبرة الناتجة عن الثقافة او المدرسة ، ويشمل ذلك القراءة والحساب وتسميه الصور التى تدل على أشخاص او اماكن معروفة (اختبار المعلومات) أو تدل على اشياء (اختبار المفردات)٠

وينعكس فى هذه البطارية التميز الشهير الذى اقترحه ريموند كاتل بين الذكاء « السائل » والذكاء « المتبلور » وخاصة فى درجتى التجهيز المتانى والتحصيل على التوالى ٠ أما التجهيز التتابعى فيدل على الكفاءة فى الانتباه والتسميع (البروفة الذهنية) وهى عمليات مستقلة عن الاستدلال والمعلومات ٠ ولا يركز كوفمان على الدرجات المنفصلة للاختبارات الفرعية الفردية (كما هو الحال فى المقياس البريطانى للقدرة) ، ولكنه يقترح تفسير أى اختلاف ملحوظ داخل كل فئة من الفئات العامة ٠

ويستخدم المقياس فى معايير الاختبارات الفرعية للتجهيز درجة معيارية معدلة متوسطها ١٠ وانحرافها المعيارى ٣ كما هو الحال فى الاختبارات الفرعية لوكسلر . أما بالنسبة للدرجات الكلية من ناحية ودرجة الاختبارات الفرعية للتحصيل من ناحية أخرى فان متوسطها ١٠٠ وانحرافها المعيارى ١٥ ٠ ويمكن للدرجات أن تتحول أيضا الى مئينيات ، والتى حسبت على المستوى القومى من ناحية وعلى اساس المستويات الاقتصادية الاجتماعية من ناحية آخرى ٠

ويمكن للطفل أن يستجيب على معظم اختبارات هذه البطارية بالاشارة أو غيرها من الاستجابات غير اللفظية ، ولا توجد مهمة فى البطارية كلها يتطلب أكثر من كلمتين كاستجابة لفظية ٠ ولهذا يقترح المؤلفان استخدام اختبار الفهم من مقياس وكسلر كمكمل لاختبارهما ٠ كما يؤكد دليل البطارية على أن الاختبارات تقيس « النشاط العقلى الراهن » وليس القدرات الفطرية٠

قياس ذكاء الرضع والأطفال فى سن ما قبل المدرسة

توجد فئة من اختبارات الذكاء الفردية اعدت خصيصا لقياس ذكاء الرضع والأطفال فى سن ما قبل المدرسة . والواقع ان قياس الذكاء فى هذه المرحلة له خصائصه ومشكلاته ، ولعل اهمها تدريب الفاحص تدريبا جيدا على أساليب الملاحظة العلمية المباشرة . وبالطبع يمكن تطبيق بعض الاختبارات الجماعية على اطفال هذه المرحلة ، الا انه بوجه عام يمكن القول ان الاختبارات الجماعية لا تصلح الا بعد وصول الطفل الى سن المدرسة (٦ سنوات) . ويترتب على ذلك أن معظم اختبارات الذكاء فى هذا السن (ان لم يكن كلها) من النوع العملى الشفوى وتتناول فى جوهرها النمو الحسى الحركى ، والخلاصة ان اختبار الاطفال فى هذه المرحلة (وبخاصة بعدما يصل الطفل الى مستوى من النمو اللغوى يسمح بالاتصال والتواصل انما هو عملية تفاعل بين الافراد inter personal من نوع رفيع . ونعرض فيما يلى بعض الاختبارات المستخدمة فى هذه المرحلة ونبدا باشهرها وهو مقياس وكسلر لذكاء أطفال ما قبل المدرسة .

مقياس وكسلر لذكاء اطفال ما قبل المدرسة :

ظهر مقياس وكسلر لذكاء أطفال ما قبل المدرسة (WPPSI) أو Wechsler Preschool and Primary Scale of Intelligence عام ١٩٦٧ ، ويهدف الى قياس ذكاء الاطفال من سن ٤ سنوات الى ٦½ سنة، ويتألف من ١١ اختبارا فرعيا يستخدم منها ١٠ اختبارات فقط فى تحديد نسبة الذكاء ، ومنها نجد ٨ اختبارات عبارة عن خفض لمستوى ، الاختبارات المناظرة لها فى مقياس الأطفال WISC آنف الذكر ، اما الاختبارات الثلاث الباقية فقد أعدت خصيصا لهذا المقياس . وتصنف الاختبارات الفرعية فى هذا المقياس - كما هو الحال فى مقياس WISC, WAIS الى قسمين : لفظى وعملى ، على النحو الآتى (وقد وضعنا علامة * امام الاختبارات الجديدة) .

المقياس اللفظى	المقياس الادائى (العملى)
المعلومات	* بيت الحيوان
المفـــردات	تكملة الصــور
الحســـاب	المتـــاهات
المتشابهات	* الرسم الهندسى
الفهم	رسم المكعبات
* الجمل (اختبار اضافى)	

ويوصى وكسلر باستخدام الاختبارات اللفظية والعملية بالتناوب لتوفير التنوع والاحتفاظ بميل الطفل وتعاونه · ويتطلب المقياس فى تطبيقه ما بين ٥٠ دقيقة و ٧٥ دقيقة فى جلسة واحدة أو جلستين اختباريتين ·

وقد حل اختبار الجمل فى هذا المقياس محل اختبار مدى (اعادة) الأرقام فى مقياس WISC ، وفيه يكرر الطفل كل جملة مباشرة بعد عرضها عليه شفويا ·وهذا الاختبار يمكن ان يستخدم كبديل لأى اختبار لفظى آخر ، أو يطبق كاختبار اضافى لتزويد الفاحص بمعلومات اكثر عن الطفل ، وفى كلتا الحالتين لا يستخدم فى حساب نسبة الذكاء · اما اختبار بيت الحيوان فيشبه اختبار رموز الأرقام فى مقياس WAIS واختبار الترميز فى مقياس WISC · ويتضمن الاختبار مفتاحا يعطى للطفل على لوحة يتألف من صور لكلب ودجاجة وسمكة وقطة ، لـــكل منها بيت (أسطوانى) ملون بلون مختلف موضوع تحت الصورة ، وعلى الطفل أن يضع الأسطوانة (البيت) فى ذات اللون المناسب فى فتحـة تحت صـورة الحيوان على اللوحة · وتتحدد درجة الطفل فى ضوء الزمن وعدد الأخطاء وعدد الاستجابات المحذوفة · أما اختبار « الرسم الهندسى » فيتطلب من الطفل محاكاة ١٠ رسوم بسيطة بقلم ملون ·

تقنين المقياس : تم تقنين هذا المقياس على عينة تتألف من ١٢٠٠ طفلا مصنفة الى ٦ فئات ، مدى كل منها نصف سنة فى الاعمار من ٤ الى ٦½ سنة وكانت كل فئة تتكون من ١٠٠ طفل و ١٠٠ طفلة اختيرت على أساس بيانات التعداد العام للسكان فى الولايات المتحدة الامريكية عام ١٩٦٠ ·

وفى تصحيح المقياس يتم تحويل الدرجات الخام الى درجات معيارية معدلة متوسطها ١٠ وانحرافها المعيارى ٣ ، اما مجموع الدرجات الموزونة فى المقياس اللفظى والمقياس العملى والمقياس الكلى فيتحول الى نسبة ذكاء انحرافية متوسطها ١٠٠ وانحرافها المعيارى ١٥ · بالاضافة الى الأعمار العقلية بالمعنى التقليدى :

وقد حسب ثبات الاختبارات الفرعية ـ فيما عدا اختبار بيت الحيوان ـ بطريقة التجزئة النصفية · أما اختبار بيت الحيوان فقد حسب ثباته بطريقة اعادة الاختبار لاعتماده الكبير على السرعة فى الاداء · وتراوحت معاملات الثبات للاختبارات الفرعية والمقاييس الكلية بين ٨٠ر ، ٩٢ر ·

أما من حيث صدق المقياس فلازال الميدان يحتاج لمزيد من البحث ، وقد يرجع ذلك الى حداثة المقياس · وقد حسبت علاقته بمقياس ستانفرد ـ بينيه فبلغ معامل الارتباط ٧٥ر فى مجموعة ٩٨ طفلا تتراوح أعمارهم بين ٥ ، ٦ سنوات · كما هو الحال فى مقياس WISC فان معامل ارتباط المقياس اللفظى بمقياس ستانفرد ـ بينيه أعلى من ارتباط المقياس العملى به· وقد أجريت عدة تحليلات عاملية للمقياس سجلها ساتلر (٧٣) وفيها تأكد وجود عامل عام بالاضافة الى عامل لفظى وعامل عملى وكان العاملان الطائفيان أقل ظهورا فى الأعمار الدنيا · وعموما فان اجراءات تقنين هذا المقياس وحساب ثباته على درجة عالية من الدقة الفنية ـ كما تقول انستازى (٤٩ : ٢٩١) ـ وهو فى ذلك يتفوق على كثير من اختبارات ذكاء أطفال ما قبل المدرسة التى ظهرت قبل ذلك ، ويعكس التطورات التى طرأت على حركة بناء الاختبارات فى السنوات الأخيرة ، ومع ذلك فان هذا المقياس لم يظهر بأية صورة فى أى بيئة عربية على عكس مقياسى وكسلر السابقين للراشدين والأطفال ·

الجداول الارتقائية لجيزل :

من المحاولات الكثيرة التى بذلت لقياس ذكاء الرضع واطفال ما قبل المدرسة ما يسمى المقاييس الارتقائية developmental scales التى تعتمد على نتائج البحث فى علم النفس الارتقائى · ومن أشهر هذه المقاييس

جداول جيزل Gesell schedules التى ظهرت عام ١٩٤٧ وتفيـــد فى تحديد مستوى النمو الذى يصل اليه الطفل (بالمعايير الأمريكية بالطبع) ٠ وتعتمد فى تقدير جوانب النمو اجراء تقنين يعتمد على الملاحظـــة المباشرة لاستجابات الطفل لمثيرات مقننة ، وتقويم سلوكه ونموه فى مواقف الحياة اليومية فالملاحظة جوهر هذه الأداة بالاضـــافة الى المعلومات التى يمكن الحصول عليها من مصادر أخرى كالأم ٠ وتشمل هذه الجداول فترة طويلة نسبيا من النمو حيث تمتد من سن ٤ أسابيع حتى سن ٥ سنوات ٠ وتؤكد نتائج البحث ان هذه الجداول تعطى معاملات ثبات بين الفاحصين المدربين يزيد على ٩٥ر٠ على الرغم من أن نظام الملاحظة والتصحيح منها اقل تقنينا اذا قورن بغيرها من الاختبارات فتدريب الملاحظين يرفع الثبات (٤٩ : ٢٥٦) ٠ وقد استخدمت هذه الجداول منذ ظهورها من جانب الاخصائيين النفسيين والاطباء العقليين سواء فى البحث العلمى أو الممارسة العملية ٠ وتشمل هذه الجداول فى صورتها الأخيرة خمســة حالات رئيسـية هى : الحركات الغليظة ، الحركات الدقيقة ، اللغة ، السلوك التوافقى ، والسلوك الشخصى الاجتماعى والتى يزود الفاحص او الملاحظ مع كل منها بوصــف لفظى منصل للسلوك المعتاد فى المراحل العمرية المختلفة بالاضافة الى رسوم توضيحية لهذا السلوك ٠

خفض مستوى مقياس ستانفرد ـ بينيه :

ادى هذا الاهتمام بذكاء الرضع وأطفال ما قبل المدرسة الى ادخال تعديلات جوهرية على مقاييس بينيه بحيث يمكن « خفض مستواها » ٠ وكانت أولى هذه المحاولات ما قام به كوهلمان Kuhlman عام ١٩٢٢ حينما عدل مقاييس بينيه الأصلية بحيث تصل الى سن ثلاثة شهور ، وكان اختباره من المحاولات المبكرة لقياس ذكاء الرضيع ثم عدل هذا المقياس عام ١٩٣٩ وظهر باسم مقياس كولمان ـ بينيه بحيث يشمل معظم فترة الطفولة المبكرة ٠

من اهم المقاييس التى استخدمت أسلوب خفض مقياس ستانفرد ـ بينيه مقياس كاتل الذى يسمى مقيــــاس كاتل لذكاء الرضيع Cattell Infant Intelligence Scale الذى يعتبره كثير من علماء النفس ملائما لقياس ذكاء الرضع ٠

وقد أعد هذا المقياس بغرض ، خفض مستوى ، تعديل عام ١٩٢٧ من مقياس ستانفرد - بينيه ، وفيه استخدم كاتل اسئلة ستانفرد - بينيه بالاضافة الى اسئلة أخرى انشئت او اقتبست من جداول جيزل وغيرها من اختبارات الرضع المتاحة ، وصنفت الأسئلة الى مستويات عمرية ، وحسب العمر العقلى ونسبة الذكاء على النحو الذى استخدم فى طبعة ١٩٣٧ من مقياس ستانفرد - بينيه ·ويشمل هذا المقياس مدى من العمر يمتد بين شهرين الى سن ٣٠ شهرا ، وخلال العام الأول تتحدد مستويات العمر فى فترات طول كل منها شهر واحد ، وفى العام الثانى تتحدد بفترات طولها شهران ، اما فى النصف الأول من العام الثالث فان طول الفترات ٣ شهور · فاذا اجتاز الطفل جميع الأسئلة المخصصة لهذه الأشهر الثلاثين الأولى من الحياة يستمر الفاحص معه بتطبيق مقيـــاس ستانفرد بـ ينيه مبتدئا بمستوى ٣ سنوات ·

وقد قنن مقياس كاتل لذكاء الرضع على مجموعة من ٢٧٤ طفلا ، اعيد اختبار كثير منهم فى الأعمار ٣ ، ٦ ، ٩ ، ١٢ ، ١٨ ، ٢٤ ، ٣٠ ، ٣٦ شهرا·

ومن الاختبارات التى تنتمى الى نفس الفئة اختبارات ميريل - بالمر Marrell Palmer واختبارات العالم الانجليزى فالانتاين C.W. Valentine والجداول الارتقائية التى أعدها فى بريطانيا العالم الانجليزى جـــــريفيت Griffith وفى رأى فرنون (٨٢ : ٦٣) ان جداول جريفيت هذه هى أكثر اختبارات ذكاء الرضع شمولا وأدقها تقنينا · فهى تتضمن ٣ أسئلة لكل أسبوع من أسابيع العمر فى العام الأول من الحياة ، وسؤالين لكل اسبوع فى العام الثانى ، وبهذا يبلغ عدد أسئلة القياس ٢٦٠ · وتصـــنف الى ٥ مجموعات ، كل منها متكافىء مع الفئات الأخرى فى الدرجات بحيث يعطى المقياس تخطيطا نفسيا (بروفيلا) للنمو فى كل فئة ، بالاضافة الى درجة كلية تسمى النسبة العامة General Quotient .

ويتضمن الجدول رقم (١٤) هذه الفئات الخمسة ، مع امثلة منها فى بداية المقياس ونهايته ·

جدول رقم (١٤) فئات مقياس جريفيث وأمثلة من أسئلته

الفئة	أمثلة من سلوك الطفل بعد الولادة	أمثلة من أسئلة المقياس عند اقتراب الطفل من سن عامين
(أ) النشاط الحركى	الرفس ، رفع الرأس	النزول على الدرج دون مساعدة أحد ٠
(ب) النشاط الاجتماعى	الابتسام ، التعرف على الأم	توجيه أسئلة اليه عن أشياء موضوعة أمامه على منضدة ٠
(حـ) السمع والكلام	الاجفال من الضوضاء	استخدام جمل تتألف من ٤ كلمات أو أكثر
(د) تآزر العين واليد	تتبع أضواء معينة	قذف كرة فى سلة ٠
(هـ) الأداء العملى	امساك الأشياء	تجميع أجزاء لعبة ٠

مقاييس بيلى لنمو الرضع :

شهد عقد الستينات والسبعينات اهتماما جديدا بالاختبارات التى تصلح للرضع وأطفال ما قبل المدرسة ، ومن أهم العوامل التى ساعدت على ذلك التوسع السريع فى البرامج التعليمية المعدة للاطفال المتخلفين عقليا من ناحية والنمو الهائل فى برامج ما قبل المدرسة فى مجال التربية التعويضية للاطفال المحرومين اجتماعيا وثقافيا وقد ظهرت اختبارات جديدة لمواجهة هذه المطالب الملحة ٠

ومن أشهر هذه الاختبارات مقاييس بيلى لنمو الرضع Bayley Scales of Infant Development ويشمل مواد من جداول جيزل ومن بعض الاختبارات الأخرى التى تصلح لهذه المرحلة٠ وتتضمن مقاييس بيلى ثلاث أدوات متكاملة لتقدير المستوى الارتقائى للطفل من سن شهرين الى ٢½ سنة وهى :

المقياس العقلى والمقياس الحركى وسجل سلوك الرضع • ويشمل المقياس العقلى عمليات الادراك والذاكرة والتعلم وحل المشكلات والتلفظ ، وبدايات التواصل اللفظى، والتفكير المجرد فى صوره المبكرة • ويشمل المقياس الحركى القدرات الحركية الغليظة ومهارات المعالجة بالأيدى والأصابع ، ويتداخل هذا المقياس مع المقياس السابق حيث يلعب النمو الحركى دورا هاما فى تفاعل الطفل مع البيئة وبالتالى فى نموه العقلى أما سجل سلوك الرضيع فيستخدمه الفاحص أو الملاحظ لتقويم الجوانب المختلفة فى نمو الشخصية مثل السلوك الانفعالى والاجتماعى ، ومدى الانتباه ، والمثابرة . والتوجه نحو الهدف••

وقد أعدت لهذه المقاييس معايير مشتقة من أداء عينة ممثلة للمجتمع الأمريكى من ١٢٦٢ طفلا موزعين بالتساوى تقريبا حسب الاعمار من شهرين الى ٣٠ شهرا واستبعدت من العينة حالات أطفال المؤسسات والأطفال المبتسرين والأطفال الذين يعيشون فى بيئات اسرية مزدوجة اللغة • ويعطى المقياس العقلى والمقياس الحركى مؤشرين منفصلين وتتحول الدرجات الخام فيهما الى درجات معيارية اعتدالية متوسطها ١٠٠ وانحرافها المعيارى ١٦ (كما هو الحال فى مقياس ستانفرد ـ بينيه) ـ وقد أعدت هذه المؤشرات الارتقائية كمعايير داخل المجموعات العمرية مصنفة على فترات طولها نصف شهر خلال الأعمار من شهرين الى ٦ شهور ، وعلى فترات طولها شهر واحد للاعمار من ٦ شهور الى ٣٠ شهرا •

وحسبت معامل الثبات لهذه المقاييس باستخدام طريقة التجزئة النصفية لكل مجموعة عمرية ، وقد بلغ وسيط هذه المعاملات للمقياس العقلى ٨٨ر ، وللمقياس الحركى ٨٤ر ، ويعطى المقياس معامل اتفاق مرتفع بين الملاحظين والفاحصين • أما الصدق فلا تتوافر عنه أدلة كافية ، ولو أن بيلى يرى أن استخدام مقاييسه أفضل ما يكون لأغراض التشخيص وليس التنبؤ •

مقاييس مكارثى لقدرات الأطفال :

من المقاييس الحديثة والهامة أيضا فى هذه المرحلة مقياس مكارثى لقدرات الأطفال McCarthy Scales of Children Abilities والذى يصلح للاطفال فى الأعمار من ٢½ الى ٨½ سنة •ويتألف من ١٨ اختبارا مصنفة

الى ستة مقاييس متداخلة هى : المقياس اللفظى والادراكى الأدائى ، والكمى، والمعرفى العام ، ومقياس الذاكرة ، ومقياس الحركة ٠

ويقترب المقياس المعرفى العام من فكرة المقياس التقليدى الكلى للنمو العقلى ، وتيم الحصول عليه من مجموع درجات المقـــاييس الثلاثة الأولى والتى تؤلف ١٥ اختبارا من ١٨ اختبارا تؤلف البطارية الكلية ٠ وتتحول الدرجات الخام فى هذا المقياس الى درجات معيارية اعتدالية متوسطها ١٠٠ وانحرافها المعيارى ١٦ حسبت لكل مجموعة عمرية مداها ٣ شهور ، اما درجات المقاييس المنفصلة فتتحول الى درجات معيارية اعتدالية متوسطها ٥٠ وانحرافها المعيارى ١٠ لنفس المجموعات العمرية ٠

وقد بلغت عينة تقنين هذا المقياس ١٠٠ طفل تقريبا لكل مجموعة عمرية من المجموعات العشر التى تمتد على فترات مداها نصف سنة للمستويات العمرية من ٢½ ــ ٥½ سنة ، ومداها عام كامل للمستويات العمـــرية من ٥½ ــ ٨½ سنة ، وبلغ قوام العينة الكلية ١٠٣٢ طفلا سويا ، تمثل المجتمع الأمريكى حسب تعداد ١٩٧٠ ٠

وحسب معامل الثبات بطريقة التجزئة النصفية فبلغ فى المتوسط ٩٣ر للمجموعات العمرية المختلفة فى المؤشر المعرفى العام ، وبالنسبة للمقاييس الخمسة الأخرى يتراوح معامل الثبات بين ٧٩ر ، ٨٨ر٠ واستخدمت ايضا طريقة اعادة الاختبار ن (= ١٢٥ طفلا) فبلغ معامل الثبات للمقيـــاس المعرفى العام ٩٠ر وتتراوح للمقاييس الخمسة الأخرى بين ٦٩ر ، ٨٩ر

أما الصدق فلا تتوافر عنه أدلة كافية باستثناء بعض البيانات عن الصدق التنبؤى باستخدام محك التحصيل المدرسى فى نهاية الصف الأول بالمدرسة الابتدائية ٠ وعموما فان البطارية تعتمد فى بنائها الأصـلى على مصادر متعددة : الخبرة الكلينيكية ونتائج البحث فى علم نفس النمو ونتائج بحوث التحليل العاملى ٠

المقاييس المعتمدة على نموذج جان بياجيه :

ظهرت فى السنوات الأخيرة مقاييس لمراحل العمر المختلفة معتمدة على

نتائج البحوث الرائدة التى قام بها عالم النفس السويسرى جان بياجيه لفترة تتجاوز نصف قرن • ومن السمات الهامة فى هذه المقاييس أنها من النوع الرتبى ordinal ، وقد عرضنا بعض خصائص هـــذا النوع من المقاييس فى الفصلين الأول والسادس ، وبايجاز نقول انها تفترض التتابع المنظم والموحد للنمو خلال المراحل المتتابعة ، وهى المراحل التى تمتد من مرحلة الرضاعة حتى مرحلة المراهقة وما بعدها ، والتى يســـميها بياجيه المرحلة الحسية الحركية ، ومرحلة ما قبل العمليات ، ومرحلة العمليات العيانية ، ومرحلة العمليات الشكلية الصورية (راجع تفصيل هذه المراحل فى كتابنا المرجع رقم ١٨) • وتتسم هذه الاختبارات ايضا بانها مرجعة الى المحتوى وتنسب الى المستوى فى التفسير وليس الى المعيار ، لأنها تقـــدم وصفا كيفيا لما يستطيع الطفل أداءه فى كل مرحلة • والمهام التى تتألف منها هذه الاختبارات تهدف الى التعرف على مفاهيم معينة لدى الطفل ، أو ما يسميه بياجيه المخططات المعرفية Cognitive Schemata (١٨) وليس قياس سمات واسعة النطاق • وبالنسبة لتطبيق هذه الاختبـــارات فانه يتطلب الحصول على تفسير الطفل لحدث معين ملاحظ وأسباب هذا التفسير ويعتمد التصحيح فى جوهره على نوعية الاستجابة وكيفهـــا وليس على عـــدد الإستجابات الصحيحة لعدد من الأسئلة ، كما هو الحـــال فى الاختبارات التقليدية • ولهذا نجد أن الاجابة الخاطئة التى تصـــدر عن الطفـــل فى الاختبارات المعتمدة على نموذج بياجيه تحتـــل اهمية خاصـــة كالاجابة الصحيحة فى فهم العمليات المعرفية لدى الطفل • وجوهر القياس هنا هو عملية حل المشكلة وليس نتائج هذا الحل •

وعلى الرغم من أن معظم الاختبارات من هذه الفئة لاتزال فى مرحلة البحث ، الا أنه توجد بعض الاختبارات تصلح للاستخدام العام نذكر منها :

١ ـ مقاييس النمو العقلى : من اعداد ليورنيـــدو Laurendeau وبينارد Pinard خلال الفترة ١٩٦٢ ـ ١٩٧٠ ضمن مشروع ضخم فى علم النفس النمو أجرى فى جامعة مونتريال بكندا ويقيس بعض مفاهيم بياجيه الأساسية وخاصة السببية (العلية) والمـــكان • وتطبـــق الاختبارات بطريقة الاستخبار الشفوى Oral questionnaires ـ كما يسميها الباحثان ـ والتى تقع فى منزلة بين طريقة بياجيه غير المقننة التى يسميها

« الطريقة الكلينيكية » والطريقة المقننة الكاملة فى الاختبارات التقليدية فجميع الأسئلة مقننة ، ولكن يمكن للفاحص أن يتخذ مسارات مختلفة بديلة من أجل مزيد من استطلاع واستكشاف عمليات التفكير لدى الطفـــل الفرد ، وذلك اعتمادا على الاستجابات المبدئية التى تصدر عنه ٠

ويتم تصحيح البروتوكول الكامل لاستجابات الطفل لكل اختبار باعتباره وحدة وذلك فى ضوء المستوى الارتقائى الذى يدل عليه كيف الاستجابات ٠ وقد أخضع الباحثان مقاييسهما للتحليل الاحصائى ، فاختارا عينة تقنين عددها ٢٥ طفلا ، ٢٥ طفلة فى كل مستوى عمرى من سن ٢ ـ ٥ سنوات ، وبلغ المجموع الكلى للعينة ٧٠٠ طفلا تمثل المجتمع الكندى المتحدث بالفرنسية فى مونتريال ٠

٢ ـ المقاييس الرتبية للنمو النفسى : أعد هـــذه المقاييس أوزجيرس Uzgiris وهنت عام ١٩٧٥ لأعمار أقل تصل الى سن شهرين وتمتد الى سن سنتين فقط ، وهى مرحلة تشمل ما يسميه بياجيه المرحلة الحسية الحركية والتى تشتمل على ٦ مراحل (راجع وصفها بالتفصيل فى المرجع رقم (١٨) ٠

وتشمل البطارية ٦ اختبارات تقيس ما يلى : ثبات الشىء ، تنمية وسائل للوصول الى غايات بيئية مرغوبة ، المباداة ، العلية الاجرائية ، العلاقات بين الأشياء فى المكان ، نمو المخططات ٠ وقد حصل الباحثان على معـــامل اتفاق بين الملاحظين ، ومعامل ثبات باعادة الاختبار ، لكل مقياس والتى تراوحت بين ٨٠٢ر ، ٩٩١ر ٠ كما حسب الباحثان الحد الأدنى لمعـــامل الرتبية Ordinality باســـتخدام طريقة جتمان المسماة Scologram فبلغ ٥٠ر٠ الا أن انستازى (٤٩ : ٢٧٤) ترى أن تطبيق هذه الطريقة على المقاييس المعتمدة على نموذج بياجيه موضع خلاف لا يتسع المقام لتناوله بالتفصيل ٠

٣ ـ مجموعة تقرير المفاهيم : وقد أعد هذه البطارية جولدشـــمد Goldschmid وبنتلر Bentler عام ١٩٦٨ للاعمار من ٤ ـ ٧ سنوات ، وتقيس أحد المفاهيم الهامة فى نموذج بياجيه وهو مفهوم ثبــات

الخصائص Conservation (راجع ١٨) ، والذى يعد مؤشرا على تحول الطفل من مرحلة التفكير قبل العملياتى الى التفكير باستخدام العمليات العيانية او المحسوسة ، وهو التحول الذى يحدد بياجيه حدوثه فى سن ٧ ـ ٨ سنوات ٠

ولهذا المقياس ثلاث صور ، منها اثنتان متكافئتان فى قياس المكان الثنائى البعد والعدد والجوهر والكم المتصل والوزن والكم المنفصل ، وقد تأكد امبيرقيا انهما صورتان متكافئتان فى المتوسط والانحراف المعيارى ، ومعامل الارتباط بينهما ٩٥ر٠ اما الصورة الثالثة فتقيس المساحة والطول ٠ ويتطلب تطبيق المقياس التزام الفاحص بتعليمات مكتوبة وتشمل عرض رسوم للمواد المستخدمة فى الاختبار وتوجيهات لمعالجتها وتناولها وتعليمات لفظية تقدم للمفحوصين ٠

وقد حسبت معايير فى صورة مئينيات من افراد عينة تقنين تتألف من ٥٠٠ طفل وطفلة تمتد اعمارهم بين ٤ ـ ٨ سنوات تمثل مدينة لوس انجليس بولاية كاليفورنيا ٠ وقد حسبت معاملات الثبات باستخدام ثبات المصححين ومعادلة كيورد ـ ريتشاردسون ، وطريقة اعادة الاختبار ، كما حسبت القياسية Scalability والرتبية للاختبار ، وبالاضافة الى وصف بنيته باستخدام التحليل العاملى ٠ الا أن العينات كانت صغيرة ، ويحتاج الاختبار لمزيد من البحث ٠ ولعل أهم النتائج التى يزودنا بها المقياس هى قابليته للاستخدام فى البحوث النفسية المقارنة أو عبر الحضارية cross-cultural.

تعليق عام على المقاييس المعتمدة على نموذج بياجيه :

على الرغم من أن الاتجاه نحو بناء مقاييس نفسية تعتمد على نموذج بياجيه قد تأخر طويلا ، الا انه يمثل فى السنوات الأخيرة أحد معالم التطور التى يشهدها علم النفس المعاصر ، الا أن ذلك يجب الا يصرفنا عن بعض المشكلات المتضمنة فى هذه المقاييس ولعل أهمها ، ظهور بعض مظاهر « عدم الاتساق » decalage فى التتابع المتوقع للنمو ، ويتضمن هذا بعض الشك فى افتراضات الاتصال والانتظام والتتابع التى يطرحها نموذج بياجيه للنمو

العقلى • فكثيرا ما يلاحظ أن المرحلة تطابق اداء الطفل تختلف من مهمة لأخرى ، ليس فقط بالنسبة للعمليات المختلفة التى تطلب للحل ، وانما ايضا حين تستخدم نفس العملية فى أنواع مختلفة من المحتوى •

الا أن ما نحب أن ننبه اليه ان هذه المقاييس ارتبطت ارتباطات مرتفعة مع اختبارات الذكاء التقليدية المقننة . وترتبط بنفس الدرجة مع التحصـيل المدرسى للاطفال فى الصف الأول ، كما هو الحال بالنسبة لاختبارات الذكاء الفردية والجماعية • ومعنى ذلك أنه على الرغم من اختلاف المنهج فان مهام بياجيه وأسئلة اختبارات الذكاء المقننة ، واسئلة الاختبارات التحصـيلية بينها جميعا قدر من التداخل والاشتراك ، وعلى الرغم من ان مهام بيـاجيه أصعب فى التطبيق وتستغرق وقتا أطول ، الا أن عائدها للتقويم النفسى كبير، فهى تقدم لنا وصفا كيفيا للنشاط العقلى للطفل لا توفره الأدوات الأخرى •

الاختبارات العملية للذكاء

توجد طائفة من المقاييس العقلية تقيس الأداء العملى ولا تتضمن أى اختبارات لفظية أو لغوية • وقد أشرنا الى أن مقاييس وكسلر تضمنت قسما عمليا أدائيا بالاضافة الى القسم اللفظى اللغوى ، وبالفعل فان المقاييس العملية للذكاء استفادت من التطورات التى حدثت فى ميدان القيـــاس النفسى بعد ظهور مقياس وكسلر ـ بليفو فى ميدان قياس الذكاء ، بالاضافة الى الجهود الى بذلت لقياس الذكاء « العملى » • ولهذا يمكن القول انه توجد فئة من اختبارات الذكاء العام تعتمد اعتمادا كليا على الأداء العملى وحده • وتتطلب نتيجة لذلك تطبيقا فرديا •

وقد يكون أول مقياس عملى للذكاء هو لوحة الأشكال التى ابتكرها سيجان Seguin عام ١٨٦٦ لتستخدم فى قياس ذكاء ضعاف العقول، وشاعت هذه اللوحة بعد ذلك ، بل وادخلت فى كثير من اختبارات الذكـاء العملية اللاحقة • وتتألف فى الوقت الحاضر من ١٠ قطع خشبية توضع أمام الطفل بترتيب مقنن ، ثم يطلب منه وضع كل قطعة فى موضعها من اللوحة بأسرع ما يمكن ، ويسمح له بثلاث محاولات لكل قطعة ، ودرجة المفحوص لكل قطعة هى أسرع زمن يضع فيه القطعة الخشبية من بين المحاولات الثلاثة•

وفى عام ١٩١٤ أعد نوكس سلسلة من الاختبارات العملية تطبق دون استخدام اللغة ، تضمنت لوحات أشكال متزايدة الصعوبة ، بالاضافة الى اختبارين أحدهما يسمى اختبار السفينة ، والثانى يسمى اختبار المكعبات ، ويتألف الأول من ١٠ قطع خشبية مستطيلة الشكل ترتب فى اطار خشبى لتكون فى النهاية صورة لسفينة فى البحر ، أما الثانى فهو اختبار للذاكرة المباشرة لسلسلة من الحركات تصدر عن الفاحص فى صورة نقر محدد تحديدا مسبقا بطريقة مقننة على كل مكعب من المكعبات الأربعة التى يتألف منها الاختبار، ثم يطلب من المفحوص أن يحاكيه ٠

وقد يكون مقياس بنتر - بيترسون Pinter-Patrson الذى ظهر عام ١٩١٧ أكثر الاختبارات أهمية من حيث التقنين ٠ وهو يتألف من مجموعة من الاختبارات مشتقة من الاختبارات العملية السابقة ، وخاصة اختبارات سيجان وهيلى ونوكس وغيرها ، ويبلغ عددها الكلى ١٥ اختبارا (وتوجد صورة مختصرة للمقياس تتألف من ١٠ اختبارات فقط) ٠

وفى عام ١٩٢٣ ظهر اختبار كوهس لرسم المكعبات Kohs Block Design Test كمقياس يصاحب مقياس ستانفرد - بينيه فى التعرف على المتخلفين عقليا او يحل محله٠ ويتألف من مجموعة من المكعبات التماثلة الملونة سطوحها الست بالألوان الآتية : الاحمر ، الأزرق ، الأصفر ، الأبيض ، والأصفر والابيض معا ، الاحمر والابيض معا ٠ وتعرض على المفحوص بطاقة - فى كل مرة - عليها رسم ملون ويطلب منه انتاجه عن طريق تجميع المكعبات المناسبة ٠ ويختلف عدد المكعبات المطلوب لمختلف الرسوم تبعا لمستويات الصعوبة من ٥ مكعبات الى ١٦ مكعبا ٠ وتعتمد الدرجة على زمن الاستجابة وصوابها ٠

ومن الاختبارات العملية للذكاء التى ظهرت فى الربع الأول من هذا القرن وذاعت شهرتها اختبارات متاهات بورتيوس Porteus التى ظهرت لأول مرة عام ١٩١٤ ، وتتألف من سلسلة من المتاهات المرسومة على الورق مرتبة حسب الصعوبة ويمكن اعطاؤها دون تعليمات لفظية وذلك باستخدام المتاهات السهلة على سبيل التمثيل ٠ وتصلح هذه الاختبارات للمستويات العمرية من ٣ سنوات حتى مستوى الرشد ، وفيها يطلب من المفحوص أن

يتتبع بالقلم الرصاص أقصر طريق من نقطة البداية الى نقطة النهاية فى متاهة الورق ، دون رفع القلم مطلقا على الورقة · وحين يقع المفحوص فى « خطأ » سواء خرج عن الطريق أو دخل فى ممر مغلق يوقفه الفاحص عن الأداء ويعطيه محاولة اخرى على متاهة مماثلة ، فاذا اخطأ فى المحاولة الثانية يعطى درجة « فشل » فى هذا المستوى · ويسمح للمفحوص فى المستويات العليا بأربع محاولات · وتعتمد الدرجة على الزمن المستغرق ، ويوضع فى الاعتبار اثناء التصحيح رقم المحاولة التى اجتاز فيها المفحوص المتاهة بنجاح · ويؤكد بورتيوس أن اختباراته تقيس قدرات الاستبصار والتخطيط التى تعجز الاختبارات اللفظية عن قياسها ·

ويمكن القول أن اكثر بطاريات اختبارات الذكاء العملية شيوعا فى الوقت الحاضر بطاريتان : أولاهما ظهرت فى الولايات المتحدة عام ١٩٣٠ وتسمى اختبارات أرثر العملية ، وثانيهما اختبارات الكسندر ·

وتتألف بطارية أرثر Arthur Point Scale of Performance Tests من أفضل الاختبارات العملية التى ظهرت قبل ذلك ، وتتكون من ٩ اختبارات هى :

١ ـ مكعبات نوكس ·
٢ ـ لوحة الأشكال لسيجان ·
٣ ـ متاهات بورتيوس ·
٤ ـ اختبار كوهس لرسم المكعبات ·
٥ ـ لوحة الأشكال التى تتألف من مربع وصليب ينقسم كل منها الى ٤ اقسام توضع فيها أجزاؤها ·
٦ ـ لوحة الأشكال المتشابهة : وهى اكثر صعوبة لأن القطع الخشبية فيها أقرب الى التشابه ·
٧ ـ المانيكان ويتألف من شكل خشبى لرجل تجمع اجزاؤه : الذراعان والساقان والرأس والجذع · والبروفيل ويتألف من قطع خشبية تجمع معا لتكون وجه انسان فى صورة بروفيل ·

٨ ـ تكملة الصور (سهل) ٠
٩ ـ تكملة الصور لهيلى Healy وهو اصعب من الاختبار السابق لأن القطع الخشبية فيه متماثلة ٠

وفى عام ١٩٤٧ ظهرت طبعة معدلة من مقياس آرثر الأدائى تتألف من ٥ اختبارات فقط تشمل تعديلات على اختبارات مكعبات نوكس ولوحة الأشكال لسيجان ومتاهات بورتيوس ، بالاضافة الى اختبار آخر لتكملة الصور مأخوذ من سلسلة هيلى ـ فرنالد ، واختبار جديد يشبه اختبار كوهس لرسم المكعبات فى أنه يطلب من المفحوص محاكاة رسوم متزايدة الصعوبة معروضة أمامه على بطاقات ، ويختلف عنه فى أن الرسوم مصنوعة من الورق المقوى وملونة بألوان عديدة ٠ ويسمى هذا الاختبار الجديد باختبار آرثر لرسوم الاستنسل ٠

وفى عام ١٩٢٦ ظهر فى بريطانيا اختبار مماثل هو مقياس كولنز ـ دريفر Collins-Drever ويتألف من ١١ اختبارا عمليا ، يضم الاختبارات التسع آنفة الذكر فى مقياس آرثر (مع اعتبار البروفيل اختبارا منفصلا ، فتصبح ١٠ اختبارات) ٠ بالاضافة الى اختبار لصورة نجمة خشبية مقسمة الى أجزاء يقوم المفحوص بتركيبها ، (وتوجد صورة اخرى له تتألف من ٧ اختبارات وتصلح للاطفال الصغار) ٠

أما الاختبار الهام الثانى فهو الاختبار الأدائى لالكسندر Alexander وهو محصلة الدراسات العاملية التى قام بها العالم الانجليزى الكسندر، ويتألف من ٣ اختبارات ، اثنان منها هما اختبارا مكعبات كوهس وبناء المكعبات اللذان اشرنا اليهما ، بعد أن عدل الكسندر اجراءات تصحيحهما بحيث تسمح باعطاء اكثر من درجة واحدة للنجاح فى رسم النماذج المختلفة وللزمن المستغرق ٠ أما الاختبار الثالث فهو اختبار جديد اطلق عليه الكسندر اسم اختبار الازاحة Passalong ، ويتألف من ٩ صناديق صغيرة كل منها يحتوى على قطعة خشبية واحدة لونها أحمر وعدة قطع لونها أزرق (هذه القطع قد تكون على شكل مربعات أو مستطيلات) ٠ وعن طريق تحريك القطع بالانزلاق ، دائريا تنتقل القطعة الحمراء من أسفل الى أعلى (دون

رفع القطـع من اماكنها مطلقا) • وقد رتبت مشـكلات الاختبار تبعا لمستويات الصعوبة ، وتصحح حسب الزمن المستغرق حتى الوصـــول الى الحـــل الصحيح •

وتحول درجات هذا المقياس الى اعمار عقلية ، وقد اعدت له معايير للذكور والاناث (الانجليز) من سن ٧ سنوات الى ١٩ سنة • وتسمى نسبة الذكاء الناتجة ، نسبة القدرة العملية Practical Ability Ratio وتصلح فى راى الكسندر فى انتقاء الطلاب للتعليم الفنى •

وبالطبع فان معظم الاختبارات التى تستخدم فى قياس ذكاء المعوقين تتخذ فى جوهرها صورة الاختبارات العملية ، ومن ذلك اختبار هيسكى ـ نبراسكا Hiskey-Nebraska لقياس ذكاء الصم الذى يتألف من ١٢ اختبارا فرعيا هى :

١ ـ انماط الخرز
٢ ـ ذاكرة الألوان
٣ ـ تعيين الصور
٤ ـ تداعى الصور
٥ ـ ثنى الورق (انماط)
٦ ـ مدى الانتباه البصرى
٧ ـ انماط المكعبات
٨ ـ اكمال الرسوم
٩ ـ الغاز المكعبات
١٠ ـ المتشابهات المصورة
١١ ـ الاستدلال المكانى
١٢ ـ ذاكرة الأرقام •

اما بالنسبة للمكفوفين قد ظهرت صور معدلة من مقياس ستانفرد ـ بينيه ومقاييس وكسلر تلائم ظروف اعاقتهم •

الفصـل الثـامن

الاختبارات الجماعية للذكاء

اهمية الاختبارات الجماعية للذكاء :

ينتمى مقياسا ستانفرد – بينيه ووكسلر الى فئة الاختبارات الفـردية للذكاء ، وهى أصلح ما تكون للاغراض الكلينيكية ، كما اشرنا فى الفصـل السابق · أما فى برامج التقويم النفسى التى تتطلب اختبار اعداد كبيرة من المفحوصين فى الميادين التربوية والصناعية والادارية والعسكرية تصبح هذه الاختبارات الفردية أكثر تكلفة فى الجهـد والمال والوقت ولذلك ظهـرت الاختبارات الجماعية لمواجهة مثل هذه الصعوبات وبخاصة فى الحـرب العالمية الأولى حين أعـد الاخصائيون النفسيون العسكريون فى الولايات المتحدة اختبارا ألفا (وهو اختبار لفظى) واختبار بيتا (وهو اختبار غير لفظى) لأغراض الاستخدام الحربى ·

خصائص الاختبارات الجماعية :

لا تتطلب الاختبارات الجماعية فى معظم الحالات فاحصا مدربا تدريبا خاصا ، كما أن الاجابات فيها تسجل على أوراق منفصلة للاجابة ، ولهذا لا تتطلب هذا النوع من العلاقات الانسانية المباشرة بين فاحص ومفحوص ·

وللاختبارات الجماعية ميزات هامة منها تيسير اجراءات التطبيق ، واتساع حجم العينات التى يمكن أن تطبق عليها هذه الاختبارات وحسن تمثيلها للاصل الاحصائى السكانى العام ، ويسر وسائل التصحيح الموضوعية وتوافر معايير أكثر دقة ، وشمول عينات التقنين التى تصل الى أكثـر من ربع مليون فى بعض الاختبارات الجماعية فى حين أنهـا لم تتعد ٤٠٠٠ مفحوص فى أحسن حالات الاختبارات الفردية ·

الا أن لهذه الاختبارات مشكلاتها أيضا ، ومن ذلك ضعف العلاقة بين

الفاحص والمفحوص ، وصعوبة تأكد الفاحص من تعاون المفحوص ، وعدم احاطته بدوافع المفحوص لأخذ الاختبار أو ميله واهتمامه به • بل يصعب على الفاحص أن يتعرف على المشكلات الطارئة على موقف الاختبار وخاصة بالنسبة للاحوال الداخلية للمفحوص أثناء الاختبار الجماعى • وكذلك فان بعض المفحوصين يواجهون صعوبات فى موقف الاختبار الجماعى ، وبالتالى يتعطل أداؤهم عنه فى موقف الاختبار الفردى • كما توجد أدلة على ان الأشخاص المضطربين انفعاليا يؤدون فى مواقف الاختبار الفردى افضل من الاختبار الجماعى • وأخيرا فان الاختبارات الفردية لا تيسر للفاحص الحصول على ملاحظات اضافية عن سلوك المفحوص أو تحديد العوامل المسئولة عن ضعف الأداء فى بعض الأسئلة او بعض الاختبارات الفرعية •

وتختلف الاختبارات الجماعية عن الاختبارات الفردية فى القالب وفى طريقة ترتيب الأسئلة • فاغلب الاختبارات الجماعية الحديثة تعتمد على أسئلة الاختيار من متعدد ، وقليل منها يستخدم أسئلة النهايات المفتوحة • كما تصنف الأسئلة المتشابهة فيها الى اختبارات فرعية منفصلة توحيدا للتعليمات • وفى كل اختبار فرعى ترتب الأسئلة عادة تبعا لمستويات الصعوبة •

وقد تنشأ بعض الصعوبات العملية من استخدام نظام الاختبارات الفرعية المنفصلة • ومن ذلك أن الفاحص الأقل خبرة والأقل دقة قد يقع فى أخطاء ضبط زمن الاختبار وخاصة فى الاختبارات ذات الزمن القصير • وحتى يمكن تحقيق قدر من التوازن بين استخدام الزمن المعد للاختبار وتنظيم الأسئلة على نحو يسمح لجميع المفحوصين بمحاولة الاجابة على جميع أنواعها بمستويات صعوبة متزايدة يلجأ بناة الاخبارات النفسية فى الوقت الحاضر الى النظام الحلزونى الشامل ـ Spiral-Omnibus وفيه تعرض أولا أسهل الأسئلة من كل نوع ثم يتلوها الأصعب من كل نوع فالأصعب، وهكذا على نسق حلزونى تصاعدى فى مستوى الصعوبة • وحتى يمكن تجنب الحاجة الى اعادة التعليمات بالنسبة لكل سؤال ولخفض عدد مرات التحول فى التأهب الناتج عن التعليمات (١٨) يلجأ بعض مؤلفى الاختبارات الى استخدام النظام الحلزونى الشامل بالنسبة لمجموعات من الأسئلة ، يتألف كل منها من ٥ ـ ١٠ أسئلة •

ويرى بعض نقاد الاختبارات الجماعية أن هذا النوع يفرض قيودا على استجابات المفحوص ، وقد وجه هذا النقد خاصة الى أسئلة الاختيار من متعدد والى بعض الأنواع الشائعة من الأسئلة مثـــل القيـــاس التمثيلى والمتشابهات والتصنيف • بل ان بعض هذه الأسئلة - فى رأى هؤلاء النقاد - تعاقب المفحوص المبتكر الذى يدرك تضمينات غير عادية فى السؤال • ويرد أنصار هذه الاختبارات على هذا النقد بأن هذه الحالات نادرة الحـــدوث وخاصة اذا كانت أسئلة الاختبار حللت تحليلا دقيقا وتوافرت عنها بيـانات صدق كافية •

وفى رأى فريق أخر من النقاد أن الاختبارات الجماعية لا توفر للفاحص فرصا لتحليل الأخطاء والبحث عن أسباب اختيار المفحوص لاجابة معينـــة، وهذا النقد صحيح بالطبع • ولهذا السبب وغيره لابد من أن يطبق مع اختباره الجماعى اختبارا فرديا على المفحوص أو أن يحصل على معلومات اضافية عنه من مصادر أخرى وخاصة فى الحالات التى تتطلب اتخاذ قرارات هامة بالنسبة له ، وخاصة فى الحالات الهامشية التى يصعب اتخاذ قرار واضح بشأنها •

وتوجد مشكلة أخرى بالنسبة لهذا النوع من الاختبارات وهى عدم المرونة التى تتمثل فى أن كل مفحوص يختبر فى جميع الأسئلة • وفى هـــذا استهلاك للوقت دون مبرر ، والأجدى أن يركز كل مفحوص على الأسئلة المناسبة لمستوى قدرته ، واذا حدث ذلك فان المفحوص يتجنب الملل الناجـــم عن العمل مع أسئلة شديدة السهولة ، أو الشعور بالاحباط والقلق النـــاجم عن محاولة الاجابة على أسئلة شديدة الصعوبة وتتجاوز مستوى قدرته • والاختبارات الفردية فى هذا الصدد لها ميزة واضحة • ولعلنا نذكر أنه فى مقياس ستانفرد - بينيه يتم انتقاء الأسئلة التى تقدم للمفحوص على أساس استجاباته السابقة ، وتعطى له الأسئلة التى تقع فى نطاق مستوى قدرته •

دور الحاسب الالكترونى فى بناء نظام اختبارى تكيفى :

بذلت فى السنوات الأخيرة محاولات للجمع بين مزايا كل من الاختبارات الفردية والجماعية ، وخاصة فى مجال المواءمة بين أسئلة الاختبار وخصائص

استجابات المفحو فرد · ويسمى هذا الأسلوب باسماء مختلفة منها النظام الاختبارى التكيفى adaptive أو التتابعى sequential أو المتفرع branched أو التفصيل tailored أو التفـريدى individulized أو المبرمج programmed أو الدينامى dynamic أو المشروط بالاستجابة response-contingent (راجع انستازى ٤٩ : ٣٠١ - ٣٠٤) · وهذه الأساليب أيسر فى التناول حين تستخدم الحاسب الالكترونى فى تطبيق الاختبار ·

وتتلخص طريقة بناء النظام التكيفى (أو غير ذلك من الصفات السابقة) للاختبار فى البدء باعطاء اختبـار قصير يسمى الاختبـار الاستكشـافى routing test تشمل أسئلته مدى واسعا من الصعوبة · وفى ضـوء استجابة المفحوص فى هذا الاختبار يوجه الى اختبار من بين عدة اختبارات ذات مستويات صعوبة مختلفة ، عادة ما يكون أطول من الاختبار الاستكشافى الا أنه يتفق مع مسـتوى قدرة المفحوص وتسمى هـذه الطريقة « طريقـة المرحلتين » ·

وتوجد طريقة أخرى لبناء هذا النظام الاختبارى تسمى « طريقة الهرم » وفيها يبدأ جميع المفحوصين بالاجابة على سؤال متوسط الصعوبة ، فاذا كانت اجابة المفحوص عليه صحيحة يوجه الى سؤال اكثر صعوبة ، اما اذا كانت اجابته خاطئة يوجه الى سؤال أسهل ، ويتكرر ذلك بعد كل استجابة على كل سؤال ·

وقد أستخدمت هاتان الطريقتان فى بناء اختبارات من نوع الورقة والقلم المعتادة أو من النوع الذى يتطلب استخدام الحاسب الالكترونى · أما الطرق الأكثر تعقيدا والتى لا تتطلب انماطا ثا مجددة مقدما من تتابع الأسـئلة فلا يمكن اجراؤها الا بالحاسب الال ·

ارية الم ثة :

تستخد براءات المعـ على الكومبيوتر فى تطبيق الاختبار نظرية الاستجابة على المفردة item response theory وهى التى تسمى نظرية

السمة الكامنة latent trait theory في بعض الاحيان . وتقـوم هذه النظرية على خاصتين أساسيتين لمنحنى المفردة أو السؤال وهما :

١ ـ مقياس القدرة : ويقدر بالعدد الكلى للاستجابات على الاختبـار . وغير ذلك من المعلومات التى تتوافر للباحث عن استجابات عينة معينة من المفحوصين على الاختبار .

٢ ـ احتمال حدوث الاستجابة الصحيحة لمفردة معينة : وهـو دالة لوضع المفحوص فى مقياس القدرة . ويحسب هذا الاحتمال من البيـانات التى تمثل نسب الذين يجتازون السؤال أو المفردة من الأفراد ذوى المستويات المختلفة فى القدرة .

ويتألف النموذج الكامل لنظرية السمة الكامنة من ثلاثة بارامترات تشتق رياضيا من البيانات الامبريقية عن مفردات الاختبار وهى :

١ ـ تمييز المفردة : ويتحدد رياضيا بميل slope المنحنى الدال على علاقة مقياس القدرة باحتمال حدوث الاستجابة الصحيحة للمفردة ، وهما اللذان يؤلفان المحور الأفقى والمحور الرأسى على التوالى فى هـذا المنحنى . ويرتبط هذا البارامتر عكسيا بالمسافة المطلوبة فى متصل مقياس القدرة حتى يزداد احتمال حدوث الاستجابة الصحيحة للسؤال . وكلمـا ازدادت قيمة التمييز يزداد الميل انحدارا .

٢ ـ صعوبة المفردة : ويتطابق هذا البارامتر مع موضع المفردة فى محور مقياس القدرة حين يكون احتمال الاستجابة الصحيحة ٥٠ر.

٣ ـ التخمين : وهذا البارامتر يستخدم فى أسئلة الاختيار من متعدد والتى من اجلها صممت خصيصا نظرية السمة الكامنة . ويمثل احتمال حدوث الاستجابة الصحيحة بمحض الصدفة ويتحدد بالخط المقارب الأدنى lower asymptote فى المنحنى الخاص بالمفردة .

ويعتمد الحاسب الالكترونى على الخصائص السابقة التى تتوافر فى مجموعة كبيرة من المفردات . فبعد أن يستجيب المفحوص على سؤال له

خصائص معينة ينتقى الكومبيوتر السؤال التالى على أساس الاستجابات السابقة التى صدرت عن المفحوص لخفض اثر اخطاء التقدير · ويمكن ايضا أن تستخدم هذه الطريقة للمواءمة بين نقطة البداية فى الاختبار أو السؤال المبدى فيه ومستوى قدرة المفحوص · وفى هذه الحالة يعتمد الحاسب الالكترونى على البيانات السابقة المتوافرة عن المفحوص مثل نتائجه فى الاختبارات الأخرى أو فى الامتحانات المدرسية أو المعلومات التى يقدمها المفحوص عن نفسه فى بداية جلسة الاختبار ·

وتعتمد درجة المفحوص فى هذه الحالة ليس على الأسئلة التى يجيب عنها اجابة صحيحة ، وانما على مستوى الصعوبة والخصائص السيكومترية الأخرى للاسئلة ·

بطاريات اختبارات الذكاء المتعددة المستويات :

تظهر مقارناتنا السابقة بين الاختبارات الجماعية من ناحية الاختبارات الفردية والاختبارات المعتمدة على الحاسب الالكترونى من ناحية اخرى أن الاختبارات الجماعية للذكاء تقدم لجميع المفحوصين نفس المجموعة من الأسئلة بصرف النظر عن استجاباتهم الفردية · ولهذا السبب نلاحظ أن أى اختبار جماعى يجب أن يشمل مدى محدود نسبيا من الصعوبة بحيث يلائم مستوى معين من العمر أو الصف الدراسى أو القدرة · وحتى يمكن توافر مقاييس للنمو العقلى لمدى واسع نسبيا قام الباحثون فى هذا الميدان ببناء بطاريات متعددة المستويات multilevel · وبهذه الطريقة يتم اختبار المفحوصين فى المستوى الملائم لهم · ويمكن للمستويات الأخرى أن تستخدم لاعادة اختبار نفس المفحوصين فى مراحل تالية ، او فى أغراض التقويم النفسى المقارن بين مختلف الأعمار أو الصفوف الدراسية · ويتوافر فى هذه البطاريات أرضية وسقف مناسبان للافراد الذى يقعون على أطراف التوزيع سواء بالنسبة للعمر أو الصف الدراسى وذلك نتيجة التداخل بين البطاريات المتتابعة · الا أننا يجب أن ننبه هنا الى أن المواءمة بين صعوبة المفردات ومستوى قدرة المفحوصين ليست بمستوى الدقة الذى يتوافر فى الاختبارات الفردية أو الاختبارات المعتمدة على الكومبيوتر ، وانما عادة ما تكون هذه المواءمة تقريبية وعادة ما تتم فى ضوء المعلومات العقلية التى

تتوافر لدى الفاحص عن المفحوص وخاصة عمره الزمنى أو صفه الدراسى وليس فى ضوء استجاباته على أسئلة الاختبار ٠

وافضـــل مجال لتطبيق الاختبارات المتعددة المستويات هو المدرسة حيث تزداد الحاجة الى المقارنة بين المفحوصين خلال سنوات عديدة ، ولهذا السبب فان معظم المستويات فى هذه البطاريات تتحدد فى ضوء الصفوف الدراسية ٠ ويتوافر فى مستويات هذه البطاريات درجة مناسبة من المتصـل سواء من حيث المحتوى أو العمليات النفسية والعقلية المقيسـة ٠ كما أن الدرجات التى يحصل عليها المفحوص من مختلف المستويات تتخذ نفس الصورة ٠ بالاضافة الى أن عينات التقنين فى هــذه البطاريات تكون فى العادة متكافئة على نحو لا يتوافر فى الاختبارات التى تقنن منفصـلة ٠ واخيرا فان بعض هذه البطاريات قد يمتد على نطاق واسع ابتداء من مرحلة الحضانة ورياض الأطفال حتى مستوى القبول فى الجامعة ٠

وتصاغ معايير معظم هذه البطاريات فى صورة نسب ذكاء انحرافية أو ما يشبهها من درجات معيارية معدلة ، وبعضها يستخدم عدة أنواع من المعايير كالمئينيات والمعايير السباعية ومعايير الفرق الدراسية ٠ وتعطى اختبارات هذه البطاريات أنواعا مختلفة من الدرجـات ومن ذلك الدرجة الكلية فى البطارية بالاضافة الى درجات منفصلة للاختبـــارات اللفظية والكمية ، أو الأقسام اللغوية وغير اللغوية فى البطارية ٠

ويختلف أسماء هذه البطاريات ٠ وترصد انستازى (٤٩ : ٣٠٥) ستة أسماء شائعة هى الذكاء والقدرة العامة والقدرة العقليـة والنضج العقلى والاستعداد الأكاديمى والقدرة المدرسية ، وهى عبارات أقرب الى الترادف فى الاصطلاح السيكومترى ، ولو أن مصـــطلح « ذكاء » يقل استخدامه أكثر من غيره فى السنوات الأخيرة ربما لأســباب ترجع الى المفاهيم المختلفة للذكاء التى عرضها كاتب هذا الفصل فى موضع سـابق (٢٤) ٠ وفى رأى أنستازى (٤٩ : ٣٠٦) أن بطاريات المستويات المتعددة تتألف من « عينة من المهارات العقلية العظمى التى تعد متطلبات أساسية للعمل المدرسى وغيره من الأنشطة الهامة فى الحياة اليومية » ، وبالتالى فهى تقيس استعداد الفرد للتعلم المدرسى ـ مثلا ـ فى كل مرحلة من مراحـل العملية التعليمية ٠

اختبارات ذكاء أطفال دور الحضانة ورياض الأطفال والصفوف الأولى من مرحلة التعليم الأساسى :

تعد دور الحضانة ورياض الأطفال ابتداء من سن ٥ سنوات والصف الأول الابتدائى ابتداء من سن ٦ سنوات أدنى مستوى عمرى يمكن اعداد اختبارات ذكاء جماعية له ، اما قبل ذلك فلا تستخدم الا الاختبارات الفردية مثل ستانفرد - بينيه • وفى سن ٥ سنوات يمكن تطبيق الاختبارات الجماعية على مجموعات من الأطفال لا تزيد عن ١٠ - ١٥ طفلا ، وفى هذه الحالة يجب على الفاحص أن ينتبه للمفحوصين كأفراد للتأكد من اتباعهم التعليمات ، ويمكن اختبار عدد أكبر من الأطفال اذا توفر أكثر من فاحص •

وفى اختبار الأطفال فى هذا المستوى يجب أن تعطى التعليمات شفويا مع التوضيح اللازم وحل التمرينات والأسئلة التمهيدية (بشرط وجود عدد كاف منها) • وعادة ما تتألف كراسة الأسئلة من صور ورسوم ، ولا تتطلب من الطفل أكثر من وضع علامة تحت الاجابة الصحيحة • وقليل منها يتطلب بعض الاستجابات التآزرية مثل رسم خطوط توصل بين نقطتين • ومعنى ذلك أن اختبارات هذا المستوى لا تتطلب من الطفل اجادة القراءة أو الكتابة. ولذلك يمكن أن تنتمى الى فئة « الاختبارات غير اللفظية non-verbal والتى يجب الا نخلط بينها وبين « الاختبارات غير اللغوية » non-language التى لا تتطلب من المفحوص معرفة باللغة سواء أكانت مكتوبة أو منطوقة ، والتى تستخدم مع الأجانب أو الأميين أو الصم • فالاختبارات غير اللفظية تتطلب من الطفل معرفة باللغة المنطوقة ، ولذلك فانها تصف محتوى الاختبار ولا تصف سلوك المفحوص • فاختبارات الفهم اللفظى يمكن أن تطبق فى هذا السن باستخدام الصور : وتقاس مفردات الطفل أو فهمه للجمل بهذه الطريقة • وتفضل انستازى (٤٩ : ٢١٩) - منعا للخلط - أن تسمى هذه الاختبارات : الاختبارات « المصورة » pictorial أو الاختبارات التى لا تعتمد على القراءة non-reading بدلا من عبارة الاختبارات غير اللفظيـــة •

ومن أمثلة الاختبارات التى تصلح لهذا المستوى اختبار القدرة العقلية لأوتس ولينون Otis-Lennon Mental Ability Test (المستوى الأولى)

الذى ظهرت طبعته الأخيرة عام ١٩٧٩ ـ ١٩٨٠ ، وله صورتان (ج ، ك) لكل المستويات التى يصلح لها هذا الاختبار (وتمتــد الى نهاية المرحلة الثانوية) ويتكون المستوى الأولى Primary level من مستويين فى الواقع : المستوى ١ (لأطفال الحضانة) والمستوى ٢ (للنصف الأول من الصف الأول الابتدائى) · وهــذان المستويان متطــابقان فى المحتــوى ، ولا يختلفان الا فى طريقة تعبير الطفل عن استجابته ، ففى مستوى الحضانة يضع الطفل دائرة حول الاجابة الصحيحة ، أما فى المستوى الثانى فان الطفل يسود بالقلم الرصاص مكان الاجابة الصحيحة فى ورقة الاجابة ·

وتعطى تعليمــــات الاختبار شفويا ، كما أن لكل سؤال زمنا محددا (حوالى ١٥ ثانية) ويتطلب الاختبار ككل من ٢٥ ـ ٣٠ دقيقة ، ويطبق على جلستين بينهما فترة راحة قصيرة ، فى كل جلســة يعطى قســـم من قسمى الاختبار · ويتكون القسم الأول من ٢٣ سؤالا من أسئلة التصنيف وفيها يطلب من الطفل أن يضع علامة تحت الصورة (أو الشكل الهندسى) الذى لا ينتمى الى فئة من ٤ صــــور (أو رســـوم هندســـية) ، أمـــا القسم الثانى فيتكون من ٣٢ ســـؤالا تقيس التصـــور الذهنى اللفظى Verbal Conceptualization ، والاستدلال الكمى quantitative reasoning والمعلومات العامة ، والقدرة على اتباع التعليمات ·

ومن أمثلة التصور الذهنى اللفظى أن يطلب من الطفل وضع علامة تحت الصورة التى تدل على نار · وفى الاستدلال الكمى يطلب منه مثلا وضع علامة تحت عدد النقط الذى يساوى عدد الأقسام المعطاة لدائرة مرسومة · وفى المعلومات العامة يطلب منه وضع علامة تحت صـــورة الشىء الذى نستخدمه فى الحديث مع الآخرين (تليفون) · وفى اختبار اتباع المعلومات يطلب من الطفل مثلا وضع علامة تحت الصورة التى تدل على زجاجة داخل مربع وعليها علامة نجمة · ويحصل المفحوص فى هـــذا الاختبار بمستوياته المختلفة على مؤشر للقدرة المدرسية ·

وقد قنن هذا المقياس فى امريكا (كما قننت مقاييس المستويات الأعلى التى يقيسها نفس الاختبار) على عينة كبيرة ممثلة للمجتمـــع الأصلى فى امريكا ، ووضعت معاييره فى صورة نسب ذكاء انحرافية، انحرافها المعيارى

١٦ • كما أن له معايير مئينية وتساعية تتحدد فى ضوء فئة العمر والصف الدراسى • وللاختبار معامل ثبـات مرتفع باستخدام طـريقة كيودر ـ ريتشاردسون وطريقة اعادة الاختبار ، كما ان له صدق تلازمى وصـدق تنبؤى جيد باستخدام درجات الاختبارات التحصيلية كمحك •

ويمكن استخدام هذه الاختبارات فى قياس قدرات الاستدلال المجرد عند الأطفال من ذوى صعوبات القراءة او التعويق التربوى من الصفوف الدراسية الأعلى •

اختبارات ذكاء تلاميذ مرحلة التعليم الأساسى :

من أشهر الاختبارات الجماعية التى تصلح ابتداء من الصف الرابع بمرحلة التعـليم الأسـاسى المسـتوى (ب) من بطـارية لورج ـ ثورنديك المتعـددة المسـتويات Lorge-Thorndike Multi-Leve Eattery والتى ظهرت عام ١٩٦٤ • ويصلح هذا المستوى ابتداء من الصف الرابع الابتدائى • ولهذا الاختبار صورتان يتضمن كل منهما ٥ اختبارات لفظية و ٣ اختبارات غير لفظية • وتشمل الاختبارات اللفظية • المفردات ، وتكملة الجمل ، والاستدلال الحسابى ، والتصـنيف اللفظى والتماثلات • أما الاختبارات غير اللفظية فتشمل : تصنيف الأشكال ، وسلاسـل الأعداد ومتماثلات الأشكال • ويتطلب الاختبار حوالى ٣٥ دقيقة للقسم اللفظى • و ٣٧ دقيقة للقسم غير اللفظى • ورغم ان لكل اختبار فرعى زمنه المحدد ، الا أن هذه الاختبارات توصف بانها من اختبارات القوة •

ومن البطاريات المتعددة المستويات التى لهـا اهمية خاصـة اختبـار القدرات المعرفية Cognitive Abilities Test من اعداد لورج وثورنديك وتتألف هذه البطـارية الكلية من مستويين أوليين (ابتـداء من مرحلة الحضـانة ورياض الأطفـال وحتى الصـف الثـالث من مرحـلة التعـليم الأسـاسى فى مصـر) ومن طبعـة متعـددة المسـتويات تشمل ابتداء من الصف الثالث من مرحلة التعليم الأساسى وحتى نهـاية المرحلة الثانوية • ويبدأ المفحوصون الذين يعطون مستويات مختلفة وينتهون منها باستخدام مجموعات مختلفة من الأسئلة ، أو بلغة البطارية « وحدات قيـاس » modules .

وتقيس الطبعة المتعددة المستويات ثمانية مستويات يتألف كل منها من عشرة اختبارات مصنفة الى ثلاث بطاريات هى :

١ ـ البطارية اللفظية وتتكون من اختبارات المفردات ، وتكملة الجمل، والتصنيف اللفظى ، والقياس التمثيلى اللفظى analogy .

٢ ـ البطارية الكمية وتتكون من اختبارات العلاقات الكمية وسلاسل الأعداد ، وبناء المعادلات ·

٣ ـ البطارية غير اللفظية وتتكون من اختبارات تصنيف الأشكال والقياس التمثيلى للاشكال ، وتركيب الأشكال ·

وقد قننت البطارية على عينة جيدة التمثيل للمجتمع الأمريكى فى المستويات الصفية العشرة المتضمنة فيها · وتتحول الدرجات الخام فيها الى مقياس موحد سى جميع المستويات · والمعايير المستخدمة فى كل بطارية هى الدرجات المعيارية الاعتدالية متوسطها ١٠٠ وانحرافها المعيارى ١٦ (أى نسبة ذكاء انحرافية) ، كما تستخدم المئينيات والتساعيات ، ويمكن الحصول على مؤشر موحد من البطاريات الثلاث · وحسبت معاملات الثبات بطريقة كيودر ـ ريتشارد سون وبلغت فى المتوسط ٩ر ، كما استخدمت طريقة الصور المتكافئة وتراوحت معاملات الارتباط بين ٧ر ، ٨ر ، اما معاملات الثبات بطريقة اعادة الاختبار فكانت اقل · وبصفة عامة فان البطارية اللفظية اكثر ثباتا من البطاريتين الأخريين ·

ومن حيث الصدق فان معاملات الارتباط بين درجات البطارية عالية وكذلك معاملات الارتباط بين الاختبارات الفرعية · وتؤكد التحليلات العاملية للبطارية وجود عامل عام فى البطاريات الثلاث قد يدل على القدرة على الاستدلال باستخدام المحتوى المجرد والرمزى · وحسب صدق الاختبار ايضا باستخدام محكات متعددة منها مقاييس المهارات الأساسية والاختبارات التحصيلية والكفاءة التعليمية ، وكانت معاملات ارتباط البطارية اللفظية كالمعتاد اعلى من البطاريتين الأخريين · وعموما فان الصدق التلازمى أعلى من صدقها التنبؤى ·

ومن الاختبارات المصرية التى تقيس الذكاء العام فى هـــذه المرحلة ما يلى :

١ ـ اختبار الذكاء الابتدائى : وهو اختبار مؤسس على اختبار بالارد Ballard الذى وضع فى انجلترا لقياس ذكاء التـــلاميذ بين سن الثامنة والرابعة عشرة من العمر · وقد نقله الى العربية اسماعيل القبانى وأدخل عليه تعديلات « قليلة » (٨ : ٨) كان من الواضح « انها ضرورية لجعله ملائما للبيئة المصرية » على حد قوله · وطبق ابتداء من عام ١٩٢٨ على نحو ٣٠٠٠ تلميذا بالمدارس « الابتدائيـــة والثانوية والأولية والورش الصناعية ومدارس المعلمين والمعلمات الأولية » ، وقام القبانى بتحليل أسئلته حتى وصل فى عام ١٩٣٠ الى الصورة النهائية للاختبار التى تتالف من ٦٤ سؤالا (بدلا من ١٠٠ سؤال فى الأصل الأجنبى) مقسمة الى قسمين : الأول يتكون من ٣١ سؤالا ، والثانى من ٣٣ سؤالا · وفى كل منهمـــا تتدرج الأسئلة حسب الصعوبة · وأسئلة الاختبار غير متجانسة ولكل منهـــا زمن خاص · وقد طبق الاختبار المعدل على عينـــة تقنين كبيرة بلغت أربعـة آلاف تلميذ من أعمار مختلفة وحسبت متوسطات الأعمار لأفراد كل عمـــر من ٧ سنوات حتى ١٥ سنة بفئات مقدارها شهرين ، واعد جدول للاعمار العقلية· وهكذا فان هذا الاختبار يستخدم مفهوم العمر العقلى ، ونســـبة الذكاء بمعناهما التقليديين عند بينيه وترمان فى النصف الأول من هذا القرن · وعلى أساس نسب الذكاء أعد القبانى جدولا لتحديد مســـتويات التفوق والتخلف العقلين على النحو الآتى :

العبقرى ١٤٠ فاكثر ·
الذكى جدا من ١٢٠ ـ ١٤٠ ·
فوق المتوسط من ١١٠ ـ ١٢٠ ·
متوسط الذكاء من ٩٠ ـ ١١٠ ·
دون المتوسط من ٨٠ ـ ٩٠ ·
الغبى جدا من ٧٠ ـ ٨٠ ·
ضعيف العقل أقل من ٧٠ ·

كما أعد جدولبديل على مستوى الذكاء بطريقة لاتحتاج الى حساب العمر

العقلى • وذلك بترتيب التلاميذ الذين اختبرهم من كل عمـــر على حسب درجاتهم فى الاختبار ، ثم قسمهم الى فئات متفاوتة فى المستوى العقلى هى ا ، ب ، جـ + ، جـ ، جـ ـ ، د ، هـ على الترتيب • والفئة (جـ) تدل على المتوسط بالطبع ، ويبين فى الجدول حدود الدرجات لكـــل طبقة من طبقات الذكاء • ولما كان تلاميذ هذه الفئات من عمر واحد تقريبا فانه كلما كـان المستوى العقلى للفئة اعلى كان ذكاء افـــرادها اكبر • وبالطبع لا تطابق « طبقات » الذكاء هذه المراتب التى تتحدد بنسب الذكاء وان كانت تقترب منها كما هو مبين بالجدول رقم (١٤) •

جدول رقم (١٤) يبين طبقات الذكاء ونسب الذكاء المقابلة لها
(٨ : ٤٣)

الطبقة	نسب الذكاء المقابلة لها	
	من	الى
هـ	٠٠	٧٨
د	٧٩	٨٨
جـ ـ	٨٩	٩٥
جـ	٩٦	١٠٤
جـ +	١٠٥	١١١
ب	١١٢	١٢١
ا	١٢٢	فما فوق

وقد حسب ثبات هذا الاختبار بطريقة اعادة الاختبار على مجموعات من التلاميذ تمثل اعمار ٩ ، ١٠ ، ١١ ، ١٢ فتراوحت المعاملات بين ٨٦ر ، ٩٠ر • اما الصدق التلازمى للاختبار فقد تحدد بمعامل الارتباط بينه وبين اختبار جوادنف لرسم الرجل فوجد أن معاملات الارتبـــاط تراوحت بين ٤٧ر ، ٧٢ر لعينات من اطفال اعمار ، ١٠ ، ١٠ ، ١١ • ويزيد مقدار هذا الارتباط كلما صغر عمر الأطفال • كما حسب معامل الارتباط بين درجات الاختبار ودرجات متاهة بورتيوس فبلغ ٥١ر لأطفال الصف الابتدائى ، ٤٧ر

لتلاميذ من عمرى ٩ ، ١٠ سنوات من فرق مختلفة ، بالاضافة الى دراسة علاقته بالنجاح المدرسى وتقديرات المدرسين للذكاء ، وكانت العينات فى كل الأحوال صغيرة العدد ٠

والواقع أن هذا الاختبار من الاختبارات الجيدة حقا ، وكان يمثل بداية موفقة لحركة القياس العقلى فى مصر ٠ وقد ظهرت طبعته الثانية عام ١٩٣٩ ، ولو أنه توبع متابعة جادة كما حدث لمقياس بينيه الأصلى فى فرنسا، أو ستانفرد ـ بينيه فى امريكا ، أو بالارد فى انجلترا لقدم لعلم النفس فى مصر خدمة جليلة ٠ وفى رأينا أنه لو توافر عليه جهد جديد لاعادة تقنينه وتعديله بعد انقضاء ما يقرب من نصف قرن على تقنينه الأخير لكان فى ذلك اسهام كبير لحركة القياس العقلى فى وطننا ٠

٢ ـ اختبار الذكاء المصور : هذا الاختبار من اعداد أحمد زكى صالح، وجرب على مستويات عمرية مختلفة وتأكد انه يصلح لقياس الذكاء ابتداء من سن الثامنة و لك أعدت معايير الاختبار (فى صورة مئينيات ونسب ذكاء تقليدية) من سن ٨ سنوات حتى ١٧ سنة ٠ وهو اختبـــار متجانس الأسئلة بمعنى انه يتألف من ٦٠ سؤالا كل منها يسأل المفحوص أن يبحث عن الشكل المخالف فى كل مجموعة من الأشكال تتكون من ٥ اشال ٠

واستخرجت معاملات الثبات بطريقة التجزئة النصفية وتحليل التباين وتراوحت بين ٧٥ر ، ٨٥ر٠ ٠

اما عن صدق الاختبار فقد ارتبط بالنتاج العقلى العام (كما يقاس باختبار القدرات العقلية الأولية) بمقدار ٣٤ر ، وتشبع فى عدد من البحوث العاملية التى أشرف عليها مؤلفه بالعامل العام تشبعات تتراوح بين ٤٨ر ، ٦١ر٠ والاختبار فى رأى كاتب هذا الفصل (٢٤) يقيس احدى قدرات الذكاء وهى التصنيف ويمكن أن يعد مقياسا لهذه القدرة وحدها ولا يتجاوز ذلك الى الذكاء بمعناه الشامل ، بالاضافة الى أن محتواه يتأثر بالعوامل الثقافية٠ وبالرغم من هذه الحدود التى تقيد استخدام هذا الاختبار كمقياس للذكاء الا انه لا يزال يستخدم فى كثير من البحوث الحديثة على أنه كذلك ، ناهيك عن قصور بيانات تقنينه التى لم تتجدد منذ أوائل السبعينات بعد وفاة مؤلفه٠

٣ ـ اختبار الذكاء غير اللفظي : هذا الاختبار من اعداد عطية محمود هنا ، عن الأصل الأجنبي للاختبـــار العقلي المتعـــدد الغير لغوي لترمان ومككال ولورج • Non Language Multi-Mental Test

ويشبه هذا الاختبار الاختبار السابق (اختبار الذكاء المصور) في انه يتطلب في جميع اسئلته ـ وعددها ٦٠ ـ ان يحدد الطفل الشكل المخالف من بين ٥ اشكال ، ومعنى ذلك انه متجانس التكوين ويقيس نفس عمليـــة التصنيف التي أشرنا اليها بالنسبة للاختبار السابق • وحسب ثباته بمعادلة كيودر ـ ريتشارد سون لفئات أعمار من ٥ر٧ سنة الى ٥ر١٤ سنة فتراوحت بين ٧٢ر ، ٨٣٠ أما صدق الاختبار فقد تحدد بمعاملات ارتباطه باختبار الذكاء الثانوي (للقباني) فبلغ ٦٥ر ، وقنن الاختبـــار على عينة من ٨٤١ طفـــلا في المرحلتين الابتدائية والاعدادية تتراوح اعمارهم بين ٦ سنوات ٩ شهور و ١٦ سنة ، ١١ شهرا •

وحسبت معايير الاختبار في صورة اعمار عقلية يمكن منها حساب نسب الذكاء بالمعنى التقليدي • وفي حالة الأطفال الذين يبلغون سن ١٤ سنة او اكثر يكون العمر الزمني في معادلة نسب الذكاء مقدارا ثابتا هو ١٤ سنة أو ١٦٨ شهرا ، وهو اجراء تجاوزه المنطق السيكومتري الحديث • وينطبق على هذا الاختبار ما قلناه عن الاختبار السابق من حيث ضيق نطاق القدرات التي يقيسها من الذكاء وعدم تحديث بيانات تقنينه •

٤ ـ اختبار الذكاء الاعدادي : يهدف هذا الاختبار الى قياس القدرة على الحكم والاستنتاج خلال ثلاثة انواع من المواقف : مواقف لفظية ومواقف عددية ومواقف تتناول الأشكال المرسومة ، (٢٠ : ٤) وهو من اعداد السيد محمد خيري الذي يرى أنه يقترب من مفهوم العامل العام عند سبيرمان ، رغم انه اختبار غير متجانس التكوين . وهو بذلك أقرب الى اختبارات الذكاء من نمط ستانفرد ـ بنيه •

وقد بدأ تقنين هذا الاختبار عام ١٩٥٨ واشتملت عينة التقنين على ٣٤١٢ تلميذا من مدارس القاهرة و ١٨٠٣ تلميذا من مدارس الوجه البحري و ٥٨٨ تلميذا من مدارس الوجه القبلي (أي أن العينة الكلية بلغت ٥٨٠٣ تلميذا) وحسب الثبات بطريقة اعادة الاختبار (ن = ٥١٤) فبلغ ٩٢ر ،

وبطريقة التجزئة النصفية فبلغ ٩١ر٠ أما فى تحديد الصدق فقد حسب معامل ارتباط الاختبار باختبار الذكاء الابتدائى (للقبانى) فوصل الى ٦٥ر٠ بالاضافة الى حساب معامل ارتباطه بتقديرات المدرسين للذكاء فبلغ ٤٠ر٠ وحسبت معايير الاختبار فى صورة أعمار عقلية ٠ وقد استخدم تقنين الاختبار فى عدد من رسائل الماجستير والدكتوراه الا أنه فى حاجة الى اعادة تقنين ٠

اختبارات ذكاء طلاب المرحلة الثانوية والجامعات :

من أشهر الاختبارات التى تقيس ذكاء طالب المرحلة الثانوية اختبار المستوى الثانى من اختبارات قدرات المرحلة الثانوية والجامعية (SCAT) أو School and College Ability Tests وله صورتان أ ، ب ٠ ويهدف هذا الاختبار الى التنبؤ بالتحصيل الأكاديمى ، ولذلك يعطى للمفحوص فيه ثلاث درجات ٠ درجة فى الاختبارات اللفظية ، ودرجة فى الاختبارات الكمية ، ودرجة كلية ٠ وتعتمد الدرجة اللفظية على اختبارين هما : فهم الجمل ومعانى الكلمات ٠ وتعتمد الدرجة العددية على اختبارين أيضا هما : العمليات العددية والمسائل العددية (حل المشكلات العددية) ٠

ويتفق هذا الاختبار مع النظرية الحديثة لقياس ما يسمى « القدرات المستثمرة » (٢٤ : ٣٤) ، ولذلك فهو يعتمد على المعرفة اللغوية والعمليات الحسابية المتعلمة فى المراحل التعليمية السابقة ٠ وهو فى هذا لا يختلف عن كثير من اختبارات الذكاء الأخرى التى تعكس مقادير متفاوتة من آثار « التعلم » ولا تقيس « الوسع الفطرى » مستقلا عن الخبرات السابقة ٠

ويمكن لدرجات الاختبارات اللفظية والكمية والكلية أن تصاغ فى ضوء مقياس مشترك يسمح بالمقارنة بينها من مستوى لآخر من المستويات التى يقيسها الاختبار ٠ كما يمكن لهذه الدرجات أن تتحول الى مئينيات أو تساعيات مشتقة من عينة تقنين شاملة لكل مستوى فيه ٠

وقد حسب ثبات الاختبار بطريقة كيودر ـ ريتشاردسون وتراوحت معاملات الثبات بين ٨٨ر ، ٩٦ر المقاييس الثلاثة ٠ وقد أجرى على هذا الاختبار تحليل دقيق للمفردات للحصول على تمييز واضح بين الفئتين

اللفظية والكمية ٠ وتؤكد النتائج تمايز الاختبار تمايزا متصاعدا مع تزايد العمر والمستوى التعليمي بما يتفق مع نظريات التنظيم العقلي ٠ اما عن صدق الاختبار فقد اهتم مؤلفوه بصدقه التنبؤي واستخدموا لهذا الغرض محكات التحصيل المدرسي ، وكانت معاملات الارتباط مع الاختبارات المدرسية أعلى منها مع التقديرات ، وكانت في المستويات الدنيا أعلى منها في المستويات العليـــا ٠

وقد اعدت مجموعة من الاختبارات العقلية لتستخدم في أغراض انتقاء طلاب الجامعات وتصنيفهم وارشادهم وتوجيههم ، ومن أشـــــهرها اختبار الاستعداد التعليمي (SAT) Scholastic Aptitude Test من اعـــداد هيئة امتحانات القبول بالجامعات في الولايات المتحدة (CEEB) College Entrance Examination Board والذي نشـــر لأول مـرة عـــام ١٩٢٦ ٠ ونظهر صور عديدة جديدة من هذا الاختبـــار كـل عام ٠ ويحصل المفحوص على درجـات منفصـــــلة للقســم اللفظي والقسم الرياضي في الاختبار ٠ ويعطى هذا الاختبار تقديرا مبدئيا لاستعداد طالب المرحلة الثانوية للعمل الجامعي ٠

ومن هذه الاختبارات أيضا امتحان برنامج القياس الجامعي American College Testing Program أو (CAT) ويتضمن ٤ اقسام هي : استخدام اللغة ٠ استخدام الرياضيات ٠ قراءة المواد الاجتماعية ، قراءة العـلوم الطبيعية ٠ وقـد أعـد هـــذا الاختبـــار لنـدكويست Lindquist ويعكس وجهة نظره في أن الامتحان يجب أن يكون عينة من العمل الجامعي. وتتداخل فيه اختبارات الاستعدادات التقليدية واختبارات التحصيل ، ويركز على المهارات العقلية الأساسية المطلوبة للاداء الناجح في الجامعات ٠

ويوجد اختبار ثالث هو الاختبار السيكولوجي لجامعة أوهايو ، أو :

(OSUPT) Ohio State University Psychological Test

وهو اختبـــار قـوة محض ، يطبق دون زمن محدد ، ويســـتغرق في العادة بين ١/٢ ساعة وساعتين ٠ كما انه في جوهره اختيار لفظي يتضمن الاضداد

والتماثلات اللفظية وفهم القراءة والعلوم الطبيعية والرياضيات والعلوم الاجتماعية والأدب ٠

ومن الاختبارات الحديثة فى الظهور اختبار التأهيل الجامعى College Qualification Test ويعطى درجات فى القسم اللفظى والعددى والمعلومات العلمية ومعلومات العلوم الاجتماعية ، والمعلومات الكلية ، والدرجة الكلية فى الاختبار ٠ وكذلك برنامج الاختبارات للكليات الأمريكية American College Testing Program (ACT) والذى يشمل اختبارات الاختيار الأكاديمى ويتألف من أربعة أقسام : الاستخدام اللغوى ، الاستخدام الرياضى ، قراءة الدراسات الاجتماعية ، قراءة العلوم الطبيعية ٠

ومن الواضح انه باستثناء اختبار SAT (الذى تكمله الاختبارات التحصيلية) فان جميع هذه الاختبارات تمثل مزيجا من الاستعدادات العامة ومعرفة مادة دراسية من المجالات الأكاديمية الأساسية ٠ وحين تستخدم الدرجات المنفصلة فان صدقها الفارق فى التنبؤ بالتحصيل فى الميادين المختلفة يصبح موضع شك ٠ ومعنى ذلك أن الدرجة الكلية هى افضل منبىء بالأداء فى جميع المقررات الجامعية تقريبا ٠ واذا كان لابد من استخدام الدرجات الجزئية فان الدرجات اللفظية هى أفضل المنبئات الفردية ، ويجب أن ننبه الى أن هذه الاختبارات لا تستخدم كبدائل لاختبارات نهاية المرحلة الثانوية (الثانوية العامة فى بلادنا) فى التنبؤ بالتحصيل الجامعى ، فالواقع أن الدراسات الحديثة (٤٩ : ٢٣٢) تؤكد أن درجات المرحلة الثانوية يمكن أن تتنبأ بالتحصيل الجامعى بنفس كفاءة هذه الاختبارات ، ان لم يكن افضل. وحين ترتبط درجات اختبارات الذكاء هذه بدرجات نهاية المرحلة الثانوية (الثانوية العامة) فان التنبؤ بالأداء الجامعى يصبح اكثر دقة ٠

ومن الاختبارات التى أعدت فى مصر لقياس ذكاء تلاميذ المرحلة الثانوية وطلاب الجامعات ما يأتى :

(١) **اختبار الذكاء الثانوى** : وقد اعده أسماعيل القبانى ويتألف من ٥٨ سؤالا تتناول سلاسل الاعداد وانتاج الجمل وادراك العلاقات بين

الكلمات وادراك السخافات والاستدلال · ويصلح للتطبيق على تلاميذ المرحلتين الاعدادية والثانوية فى المدى العمرى من ١٢ سنة الى ١٨ سنة وعلى الكبار ممن تزيد أعمارهم على ١٨ سنة · وقد أعدت معايير الاختبار فى صورة طبقات للذكاء هى : المستوى (أ) وهو أعلى المستويات ، ثم المستوى (ب ، جـ ، د) ثم المستوى (هـ) وهو أدنى المستويات ، كما أعد جدول يتضمن الدرجات التى تحدد كل طبقة من كل مستوى عمرى · ولا تتوافر لدينا بيانات سيكومترية عن هذا الاختبار ·

(٢) اختبار الذكاء العالى : من اعداد السيد محمد خيرى على أساس مفهوم سبيرمان للعامل العام على أنه ادراك العلاقات والمتعلقات ويتشابه مع اختبار الذكاء الاعدادى فى أنه يقيس القدرة على الحكم والاستنتاج خلال ثلاثة أنواع من المواقف : لفظية وعددية ومواقف تتناول الأشكال المرسومة· وقد قنن على عينة عددها ٥٨٢٨ طالبا فى المدارس الثانوية والمعاهد العليا والجامعات · واشتملت العينة على المستوى الثانوى بفرعيه العلمى والأدبى، كما اشتملت عينة المستوى الجامعى على كليات نظرية وأخرى عملية ·

وحسب ثبات الاختبار بطريقتى اعادة الاختبار والتجزئة النصفية فبلغ معامل الثبات ٨٤٥ر ، ٨٨١ر على التوالى · أما فى تقدير الصدق فقد حسب معامل ارتباطه باختبار الذكاء الثانوى فبلغ ٦٩٤ر ، كما حسب ارتباطه بتقديرات المدرسين لذكاء تلاميذهم فى التعليم الثانوى فبلغ ٥٢٢ر· ، وأخيرا حسب ارتباطه بدرجات أفراد العينة الذين فى السنة الاعدادية أو الأولى من المستوى العالى أو الجامعى فى نهاية العام فكان ٥١٨ر· وأعدت للاختبار معايير فى صورة مئينيات · وهذا الاختبار حاله كحال سابقه اختبار الذكاء الاعدادى يحتاج الى اعادة تقنين ·

(٣) اختبار الاستعداد العقلى للمرحلة الثانوية والجامعات : هذا الاختبار من اعداد رمزية الغريب وينقسم الى ٥ اقسام هى :

اليقظة الذهنية · التفكير الرياضى ، ادراك العلاقات المكانية ، فهم الرموز اللغوية ، التفكير المنطقى · وهو من اختبارات القوة ، ويحتاج فى المتوسط الى ساعة ونصف للانتهاء منه · وقد حسب الثبات بطريقة

التجزئة النصفية (ن = ٥٥٠) فبلغ معامل الارتباط ٩٢ر٠ واستخدمت طريقتان لحساب الصدق التلازمي هما حساب معامل ارتباطه باختبار للقدرات العقلية الأولية فبلغ ٧٧ر٠ ، والمقارنة بين المجموعات المتضادة وكانت الفروق بين الممتازين عقليا ، والمتخلفين عقليا لها دلالة احصائية ٠

اختبارات طلاب الدراسات العليا :

يمكن القول أن الاختبارات التي تستخدم في قياس طلاب الدراسات العليا تمثل مزيجا من اختبارات الذكاء العام واختبارات التحصيل ٠ ومن اشهر هذه الاختبارات امتحانات سجل الدراسات العليا (GRE) Graduate Record Examinations والتي تطبق على الطلاب المتقدمين للالتحاق بالدراسات العليا في أمريكا ٠ وقد ظهرت هذه السلسلة من الاختبارات عام ١٩٣٦ بهدف توجيه الطلاب وقياس كفاءة التدريس وانتقاء افضل العناصر للحصول على المنح والبعثات والتعيين في وظيفة معيد ٠

ويتكون الاختبار في احدى صورة من اختبار للاستعداد واختبار تحصيلي في ميدان تخصص الطالب من مستوى رفيع مثل الأحياء والسياسة والفلسفة وعلم النفس واللغة الأسبانية ، الخ ٠ والاختبار في جوهره اختبار استعداد دراسي أو ذكاء عام ملائم لطلاب الدراسات العليا في بداية التحاقهم بها ٠

وللاختبار صورة أخرى لا تتألف فقط من اختبارات استعدادات وتحصيل رفيع ، وانما تشمل كذلك ثلاثة اختبارات من نوع اختبارات «المجال» هي العلوم الاجتماعية والانسانيات والعلوم الطبيعية ٠ وهذه الاختبارات تؤكد فهم المفاهيم الأساسية والقدرة على تطبيق المعلومات وليس مجرد استدعاء الحقائق التفصيلية ٠

وتحول درجات هذا الاختبار الى درجة معيارية معدلة متوسطها ٥٠٠ وانحرافها المعياري ١٠٠ ٠ وحسب له الثبات بطريقة كيورد - ريتشاردسون وطريقة التجزئة النصفية ، وتتراوح معاملات الثبات بين ٨٤ر ، ٩٥ر اما الثبات بطريقة اعادة الاختبار فيتراوح بين ٨٥ر ، ٩٧ر بفاصل زمني مقداره

عام واحد ، وبين ٧١ر ، ٩٢ر بفاصل زمنى مقداره عامان · وحسب الصدق التنبؤى فى ضوء محكات معينة مثل الدرجات الأكاديمية فى الدراسة العليا، والنجاح فى مقابل الفشل فى الدراسة ، وتقديرات الاساتذة ، واداء الطلاب فى امتحانات التاهيل للدكتوراه ·

اختبارات ذكاء الراشدين المتفوقين :

يمكن القول ان اختبارات طلاب الجامعة والخريجين وطلاب الدراسات العليا تلائم مستوى الراشدين المتفوقين لأغراض التقويم والبحث فى الميدان المهنى · ومن ذلك أن اختبار ميللر للقياس التمثيلىMiller Analogies Test وهو يصلح لتقويم طلاب الدراسات العليا ، يمكن ان يستخدم فى هذا الغرض· ويتألف من متماثلات مركبة مصاغة فى محتوى اكاديمى وعلى درجة كبيرة من الصعوبة ، مما يجعله ينتمى الى فئة اختبارات القوة · ومن هذه الاختبارات أيضا اختبار اتقان المفاهيم Concept Mastery Testوقد ظهر نتيجة البحوث الطولية الرائدة التى قام بها ترمان فى ميدان المتفوقين ، ويتألف من اسئلة القياس التمثيلي والمتشابهات والأضداد التى تنتمى الى محتوى متنوع فى مجالات العلوم الفيزيائية والبيولوجية والرياضيات والتـــاريخ والموسيقى وغيرها ·

اختبارات الذكاء المتحررة من اثر الثقافة :

تسعى هذه الاختبارات الى التحكم فى بعض الأبعاد التى تتألف منها الثقافة Culture ، والتى تختلف فيها الثقافات ـ وعلى وجه الخصوص بعد اللغة · فاذا كان المفحوصون يتحدثون لغات مختلفة يجب أن يتحــــرر الاختبار من هذا العنصر سواء من جانب الفاحص او المفحوص · اما حين تختلف الأطر التعليمية اختلافات واسعة وخاصة فى المجتمعات التى تنتشر فيها الأمية ، كأغلب الدول النامية ـ فان الاختبارات يجب ان تتحرر من عنصر القراءة ـ كما اشرنا الى ذلك من قبل · وفى حالة الاختبارات المتحررة من اثر القراءة Non-Reading لا يستبعد اثر اللغة « الشفوية » بالطبع ، لان هذه الاخبارات تعد لجماعات تتحدث بنفس اللغة · ومن العوامل الأخرى التى تختلف فيها الثقافات عامل السرعة speed ، وتتفاوت فى ذلك ابتداء من الايقاع الشخصى personal tempo وحتى الدافعية للعمل بسرعة

وتقدير العمل السريع ، وفى هذا تختلف الثقافات الفرعية للمجتمع الواحد (وخاصة بين الريف والمدينة) • واخيرا فان الثقافات المختلفة (فرعية أو عامة) تختلف فى التعامل مع محتوى الاختبارات • وقد حاول بعض المؤلفين اعداد اختبارات غير لغوية وغير قرائية ، ولكنها تتألف من صور لأشياء واقعية كالأثاث والألعاب والآلات (راجع فى ذلك اختبارات المقياس العملى لرتسلر ، واختبارات الذكاء المصور والذكاء غير اللفظى) وبالطبع فان هذا المحتوى « السيمانتى » للصور يعتمد على اطار الخبرة السابقة •

ولهذه الأسباب وغيرها — ظهرت اختبارات الذكاء المتحررة من أثر الثقافة Culture-free أو الاختبارات التى تصلح لأغراض المقارنة بين الثقافات Cross cultural وكان اول اختبار ظهر فى الميدان اختبار بيتا الحربى أثناء الحرب العالمية الأولى . وفيه تم التحكم فى العنصر اللغوى سواء من جانب الفاحص أو المفحوص • وكانت تعطى التعليمات بالاشارة والايماءات والتمثيل ، أما استجابات المفحوص فكانت تتمثل فى مجرد وضع علامات بسيطة (فيما عدا اختبار واحد كان يتطلب من المفحوص كتابة الأعداد) • وقد أعيد تقنين هذا الاختبار — للاغراض غير العسكرية — عام ١٩٤٦ وأصبح يتألف من ٦ اختبارات فرعية هى : المتاهات ، وتعويض الرموز والأعداد ، والسخافات المصورة ، ولوحة الأشكال المصورة ، وتكميل الصور ، والسرعة الادراكية • ويتطلب فى هذا التعديل بعض التعليمات اللفظية الا أنه يعتمد اعتمادا كبيرا على الأسئلة التدريبية •

وقد ظهرت عدة اختبارات فى هذا الميدان منها اختبارات سبيرمان الحسية للذكاء التى أعدها باللغة العربية اسماعيل القبانى وعبد العزيز القوصى ، المقياس الأدائى العالمى لليتر Leiter International Performance وسوف نقتصر على ٣ اختبارات هامة هى : اختبارات كاتل المتحررة من اثر الثقافة . واختبار المصفوفات المتتابعة لرافن واختبار رسم الرجل لجود انف •

(١) اختبارات كاتل غير المتحيزة للثقافة : أعد هذه الاختبارات العالم الأمريكى المعاصر ريموند كاتل ، ولها ثلاثة مستويات : المقياس الأول للاعمار من ٤ سنوات الى ٨ سنوات والراشدين المتخلفين عقليا ،

والمقياس الثانى للاعمار من ٨ سنوات الى ١٣ سنة والراشدين العاديين ، والمقياس الثالث من سن ١٣ سنة الى سن ١٩ سنه والراشدين المتفوقين · ولكل مقياس من هذه المقاييس الثلاث صورتان متكافئتان · ويتطلب المقياس الأول تطبيقا فرديا أو جماعيا · ويتألف المقياس الأول من ٨ اختبارات ، اربعة منها فقط تعد من نوع الاختبارات المتحررة ثقافيا · أما الاختبارات الأربعة الأخرى فتتطلب كلا من الفهم اللفظى والمعلومات الثقافية الخاصة · امـا المقياسان الثانى والثالث فمتشابهان فى المحتوى ومختلفان فى مســتوى الصعوبة ، ويتألف كل منهما من اربعة اختبارات هى : سلاسل الأشـكال ، وتصنيف الأشكال ، ومصفوفات الأشكال ، والشروط · وتتوافر للمقياس الأول فقط معايير نسب ذكاء تقليدية ، أما المقياسان الثانى والثالث فمعاييرهما من نوع نسب الذكاء الانحرافية بانحراف معيارى ٢٤ أو ١٦ · وقد شاع التحويل الاخير حتى يسمح للمقاييس بالمقارنة بغيرها ·

وقد نقل المقياس الثانى الى اللغة العربية احمد عبد العزيز ســـلامة وعبد السلام عبد الغفار ، ولا تتوافر عنه ـ فى حدود علمنا بيانات تقنين ، أما المقياس الثالث فقام باعداده فؤاد أبو حطب وآمال احمد مختار صادق وقام بتقنينه عدد من طلاب الدراسات العليا فى مصر فى بحـــوثهم للماجستير والدكتوراه · كما استخدمت آمال صـادق المقياسـين الثـانى والثالث فى دراستين عامليتين للقدرة الموسيقية (٩) وأكدت النتائج تشـبع المقياسين بالعامل العام فى عينتين مصريتين احداها من الأطفال والأخرى من المراهقين والراشدين · وقياس هذه الاختبارات للعامل العـام هـو الغرض الأساسى الذى تقوم عليه اختبارات كاتل متأثرا فى ذلك بنظرية أستاذه سبيرمان ·

(٢) **اختبارات المصفوفات المتتابعة لرافن :** ظهر هـذا الاختبار فى انجلترا لقياس العامل العام عند سبيرمان ايضا ، وتتطلب فى جوهرها ادراك العلاقات بين وحدات مجردة ، وهو فى رأى معظم علماء النفس فى بريطانيا أفضل مقاييس العامل العام · ويتألف الاختبار من ٦٠ مصفوفة (أو رسما) حذف منها أحد اجزائها ، وعلى المفحوص ان يحدد الجزء المحذوف بالانتقاء من بين ٦ أو ٨ بدائل · وتتجمع الأسئلة فى ٥ سلاسل ، يحتوى كل منها على ١٢ مصفوفة متزايدة الصعوبة ولكنها متشابهة من حيث المبدأ · وتتطلب المصفوفات الأولى دقة فى التمييز · أما المصفوفات المتأخرة والأكثر صعوبة

فتتضمن ادراك العلاقات المنطقية · وليس لهذا الاختبار زمن محدد (أى ينتمى الى فئة اختبارات القوة) · ويمكن تطبيقه فرديا أو جماعيا ·

وقد نقلت هذا الاختبار الى اللغة العربية الهيئة الفنية للخدمة السيكولوجية العسكرية ، ولا تتوافر لدينا بيانات عن خصائصه السيكومترية للعسكريين، كما كتب الدارس العراقى قيس عبد الفتاح مهدى الى احد مؤلفى هذا الكتاب (فؤاد أبو حطب) حول الدراسات التى اجريت حول هذا الاختبار بمركز البحوث التربوية والنفسية بجامعة بغداد وخاصة المصفوفات الملونة الذى تجرى محاولة تقنينه على المرحلة الابتدائية ·

وقد تم تقنين الاختبار على البيئة السعودية عام ١٩٧٧ بواسطة فريق من الباحثين باشراف فؤاد أبو حطب (٢٨) حيث طبق الاختبار على عينة تتالف من ٤٩٣٢ مفحوصا من البنين والبنات تمتد أعمارهم من ٨ سنوات وحتى ما بعد سن الثلاثين اختيروا من تلاميذ المرحلة الابتدائية والمتوسطة وطلاب المرحلة الثانوية والمرحلة الجامعية والخريجين بالمنطقة الغربية بالمملكة العربية السعودية · وحسب صدق التكوين الفرضى للاختبار ، حيث اكد صدق المحتوى أنه اختبار غير لغوى ، كما أكدت معاملات الارتباط بين درجات الاختبار والزمن المستغرق فى أدائه أنه اختبار قوة (حيث لم تكن هذه المعاملات دالة) · وثبتت أيضا صحة الترتيب التتابعى للمجموعات الخمس لأسئلة الاختبار باعتبارها متدرجة فى الصعوبة وذلك من خلال المقارنة بين متوسطات عينات الأعمار المختلفة فى كل مجموعة من المجموعات الخمس للاسئلة بالنسبة لكل درجة كلية يحصل عليها المفحوص · ثم تأكد أيضا أن تتابع المفردات داخل كل مجموعة من المجموعات الخمس منتظم من السهل الى الصعب وذلك من خلال فحص معاملات سهولة المفردات فى الأعمار المختلفة · كما استخدم فى حساب صدق التكوين الفرضى طريقة تمايز العمر وتمايز الصفوف الدراسية ، وقدر كذلك الصدق المرتبط بالمحكات باستخدام بعض الاختبارات العقلية الأخرى والتحصيل الدراسى وكانت معاملات الارتباط دالة ·

وحسب الثبات بطريقة كيودر ـ ريتشاردسون وطريقة اعادة الاختبار وكانت المعاملات مرتفعة · وحسبت المعايير المئينية للاختبار ·

(٣) اختبار رسم الرجل : اعدت هذا الاختبار الباحثة الأمريكية جود انف Goodenough عام ١٩٢٦ ، ويتطلب من المفحوص باختصار «رسم صورة لرجل» . وقد ظل هذا الاختبار فى ميدان الاستخدام العام دون تعديل منذ ظهوره حتى عام ١٩٦٣ حيث ظهر تعديل جديد شامل له باسم «اختبار الرسم لجودانف وهاريس Goodenough-Harris Drawing Test . ويؤكد هذا التعديل كما يؤكد الاختبار الأصلى ـ دقة الطفل فى الملاحظة وارتقاء تفكيره المجرد ، وليس المهارة الفنية فى الرسم ، حيث تعطى درجة لكل جزء من الجسم رسمه الطفل ، وتفاصيل الملبس ، والنسب ، والمنظور ، وغير ذلك من الخصائص التى بلغت فى طبعة ١٩٦٣ (٧٣ وحدة قابلة للقياس) . وقد اختيرت هذه الوحدات على اساس تمايز العمر ، والعلاقة بالدرجة الكلية فى الاختبار ، والارتباط بدرجات اختبارات الذكاء الجماعية الأخرى . وتم الحصول على ذلك من اختبار ٥٠ طفلا ، و ٥٠ طفلة فى كل مستوى دراسى ابتداء من الحضانة حتى سن ١٥ سنة . وقد أجريت دراسات كثيرة على المقياس المعدل ثبت منها انه ثابت وصادق ، وانه اختبار غير لفظى للنضج العقلى بالاضافة الى حساسيته الشديدة للعوامل الثقافية التى تؤثر فى تمثيل الطفل لمفهوم شكل الانسان ، ومعنى ذلك أن هذا الاختبار ليس متحررا من اثر الثقافة كما كان مفترضا .

ويتطلب المقياس المعدل ان يرسم الطفل ايضا صورة لامراة ، ويصحح هذا المقياس الجديد الخاص برسم المراة فى ضوء ٧١ وحدة تتشابه مع مقياس رسم الرجل . كما يطلب من الطفل ايضا ان يرسم نفسه Self Scale ولو ان هذا المقياس الثالث أعد فى جوهره كأسلوب اسقاطى لدراسة الشخصية ، والبيانات المتوافرة عنه ليست مقنعة . واعدت معايير لمقياس رسم الرجل ومقياس رسم المراة من تطبيقه فى امريكا على ٣٠٠ طفل من كل مستوى عمرى ابتداء من سن ٥ سنوات حتى سن ١٥ سنة . ويمكن تحويل الدرجات الخام فى كل مقياس الى درجات معيارية معدلة متوسطها ١٠٠ وانحرافها ١٥ ، وحسب له الثبات بطريقة اعادة الاختبار وطريقة التجزئة النصفية . واستخدم فى حساب صدقه محك تمايز العمر ومحك التنبؤ بالنجاح المدرسى ، كما استخدم فى بعض الدراسات مقياس ستانفرد ـ بينيه كمحك أيضا ، وكذلك تقديرات المعلمين وتتابعت الدراسات السيكومترية حول الاختبار فى مختلف الثقافات تاكدت الدراسات الأمريكية اللاحقة جميعا ثبات

الاختبار وصدقه · وقد قنن الاختبار فى البيئة المصرية وعدد من البيئات العربية الأخرى ·

وقد تكون أولى الدراسات التى أجريت فى هذا الصدد دراسة باشراف القوصى (١٤٦) والتى قام فيها ميشيل اسكندر (١٤٦) بين عامى ١٩٢٤ ، ١٩٢٧ بتطبيق الاختبار على عينة مصرية اقتربت من الألفين ، أعمارهم بين ٤ ، ١٢ سنة ، ووضع نموذجا وسطا لكل سن كما أعطى معيارا مصريا للعمر العقلى ، وتحدد صدق القياس باستخدام محك اختبار الذكاء الابتدائى للقبانى ، كما تمت المقارنة بين الصفوف الدراسية وبين الأعمار وبين الأطفال العاديين والمتخلفين وفى جميع الحالات كان التمييز فى الاتجاه المتوقع على نحو يدل على صدق التكوين الفرضى · ولم يشر البحث الى حساب الثبات ، وبعد ذلك قام على علوى شلتوت بتعديل طريقة تصحيح الاختبار بما يلائم البيئة المصرية ، وظل الاختبار يستعمل فى العيادة النفسية بكلية التربية جامعة عين شمس منذ ذلك الحين ·

وقد قام مصطفى فهمى (٤٥) بدراسة على هذا الاختبار حيث طبق على٩٧٦ تلميذا وتلميذة فى فئاتعمرية من ٦ سنواتحتى ١٢ سنة، حسبمن درجاتها معايير الأعمار العقلية مستخدما وحدة عمرية طولها سنة كاملة · ويلاحظ على هذا البحث بعض المشكلات المنهجية منها استخدام معادلة غير واضحة لحساب ثباته بطريقة كيودر ـ ريتشادسون ، فقد أشار فى مراجعه الى معادلة معينة واستخدم فيما يبدو معادلة أخرى ، بالاضافة الى الاعتماد على ٤٠ مفردة فقط ، بينما العدد الأصلى للمفردات ـ كما حددته جود انف فى الاختبار الأصلى الذى اعتمد عليه فى هذا البحث ـ هو ٥١ · كما لوحظ عدم وضوح محك تمايز العمر المستخدم فى تقدير الصدق ·

ومن الدراسات الهامة التى أجريت على هذا الاختبار تلك التى قام بها الباحث السودانى مالك البدرى (١٣٨ أ فى ٣٧) وفيها لخص البحوث العربية السابقة التى أجريت حول الاختبار ومنها رسالة الماجستير التى تقدم بها محمد بطانية الى الجامعة الأمريكية فى بيروت ، حيث طبق الاختبار على ٥٠٠ طفل أردنى تراوحت أعمارهم بين ٥ و ١٠ سنوات · وحسب الثبات بطريقة اعادة الاختبار (ن = ٣٠) فبلغ ٨١ر٠ واستخدم طريقة تمايز العمر فى

تقدير الصدق فوجد أن الفروق بين متوسطات الأطفال فى مختلف الأعمار فروق دالة احصائيا وفى اتجاه الزيادة مع التقدم فى العمر · ولكن حين حسب معامل الارتباط بين درجات الاختبار وتقديرات المستوى الدراسى لم تكن لهذا المعامل دلالة احصائية واعتبر الباحث أن سبب هذه النتيجة اما عدم دقة المعلمين فى تقدير مستوى التلاميذ أو عدم اعتماد التحصيل على الذكاء وحده · وغاب عن الباحث تفسير ثالث قد يكون أكثر صحة وهو عدم ارتباط الاختبار بمحك التحصيل · وهى نتيجة تتفق مع معظم النتائج التى توصل اليها الباحثون فى عدد من الثقافات الأخرى · واستخرج محمد بطانيه معايير مؤقتة لذكاء الطفل الأردنى ·

وفى العراق قام عبد الجليل الزوبعى (٣٧ : ٦٥) بتطبيق الاختبار على عدد من الأطفال وقارن بين نسب الذكاء حسب تعليمات جودانف والتحصيل المدرسى فلم يجد مرة أخرى ارتباطا عاليا ، واستنتج الباحث من ذلك الاستنتاج الذى أشرنا اليه فى الفقرة السابقة ، وهو احتمال أن هذا الاختبار لا يصلح لقياس ذكاء الطفل العراقى ·

وفى السودان قام مالك البدرى (٣٧) بتطبيق الاختبار على ١٤٣٥ طفلا وطفلة من سن ما قبل المدرسة والمرحلة الابتدائية فى الريف والحضر السودانيين ، وصحح الاختبار تبعا لمفتاح جودانف الأصلى (٥١ مفردة) · ولتقدير الصدق حسب معامل الارتباط بين الاختبار والمستوى الدراسى فلم يجد الارتباط دالا احصائيا مرة ثالثة · وفى ذلك يقول الباحث بأن الاختبار بصورته الحالية ، قد لا يكون صالحا لقياس ذكاء الطفل العربى والافريقى· وأن العمل الأساسى للاخصائيين النفسيين العرب فى هذا المجال ليس هو تطبيق الاختبار واستخراج المعايير فحسب ، بل اعادة النظر فى اختبار رسم الرجل بشكله الحالى ومحاولة تعديله حتى يلائم الطفل العربى والافريقى فى بيئاته المختلفة ، (٣٧ : ٧٠) · وقد استطلع الباحث مجموعة من الفروض حول صدق الاختبار قد يكون أهمها أن من أسباب اعاقة تطبيق اختبار رسم الرجل فى البيئات العربية والافريقية الاختلافات الكبيرة لأزياء الرجل لو قورن بزى الرجل الأوربى والأمريكى الأكثر ثباتا · وتأكدت صحة هذا الفرض من مقارنة متوسط درجات من يرسمون رجلا بالملابس الأوربية بمتوسط من يرسمون رجلا بالملابس التقليدية · ولذلك قام الباحث بفحص

مجموعة من رسوم الأطفال الذين يرسمون الرجل بالملابس التقليدية ، واقترح بعض التعديلات الجوهرية على مفردات الاختبار الأصلى واضافة بعض العناصر الجديدة لاعطاء من يرسمون الرجل بالملابس التقليدية فرصة متساوية مع اولئك الذين يرسمون رجلا بالملابس الأوربية • ومن اهم العناصر التى اقترح الباحث اضافتها : رسم الشارب أو اللحية أو كليهما ، اظهار الوشم، تظليل الجسم ، رسم العصا فى اليد • ويبدو ان نتائج مالك البدرى تصلح لكثير من الدول العربية التى يرتدى الرجال فيها « الجلابية » أو «الدشداشة» أو « الثوب » كالسعودية والكويت ومنطقة الخليج وبعض انحاء مصر والمغرب العربى •

وفى الكويت أجرى محمد نسيم رافت (١٤٤) دراسة على الاختبار تهدف الى تحديد نوعين من المعايير احدهما يناسب الزى الأوربى وثانيهما الزى التقليدى الكويتى وقد اجريت الدراسة على ٢٥٦٢ طفلا وطفلة تمتد اعمارهم بين ٤ ، ١٢ سنة •

وفى لبنان قام عبد السلام عبد الغفار والسيد فؤاد اعظمى بتطبيق الاختبار على عينة مكونة من ١٨٥٥ طفلا وطفلة تتراوح اعمارهم بين ٦ ، ١٢ سنة واستخدم فى تصحيح الاختبار ٥٠ مفردة فقط من مفتاح جودانف الأصلى ، وحسب الثبات بطريقة اعادة الاختبار فبلغ الارتباط ٧٥ر، وفى حساب الصدق استخدم محك تمايز العمر ، واعدت معايير اعمار عقلية•

وفى اليمن قام فتحى السيد عبد الرحيم (١٢٩) بتطبيق الاختبار على اطفال منطقة صنعاء من ذوى الأعمار الزمنية التى تمتد من ٦ ــ ١٢ سنة بلغ عددهم ١٤٨٢ طفلا وطفلة • وحذا الباحث فى بحثه حذو بحث مصطفى فهمى السابق الاشارة اليه ، فاستخدم مفرداته الأربعين وحسب معايير العمر لوحدة زمنية طولها عام كامل واستخدم طريقة كيودر ــ ريتشاردسون فى حساب الثبات • وفى حساب الصدق استخدم محك تمايز فكانت الفروق دالة بين الأعمار ، كما حسب معامل الارتباط بين الاختبار ومحك تقديرات المعلمين فبلغ ٥٩ر وهو معامل مرتفع نسبيا بالمقارنة بالمعاملات التى لوحظت فى البحوث السابق الاشارة اليها •

وقد اجريت هذه الدراسات على اختبار جودانف الأصلى طبعة ١٩٢٦، اما بالنسبة للاختبار المعدل لجودانف وهاريس الصادر عام ١٩٦٣ فلم يجر عليه الا عدد محدود من الدراسات فى الثقافتين المصرية والعربية ٠ فقد قام محمد فرغلى فراج وعبد الحليم محمود السيد وصفية مجدى بترجمة كراسة تعليمات هذا الاختبار عام ١٩٧٦ ، وفى نفس العام اجرى محمد متولى غنيمه (٤٤) تقنين هذا المقياس الجديد على البيئة المصرية حيث طبق الاختبار على ١٩١٥ طفلا وطفلة من سبع مدارس ابتدائية بمحافظة القاهرة ٠ وعلى الرغم من أن بعض الأطفال رسموا الرجل مرتديا الجلباب الا أن الباحث اعتمد فى التحليل على أولئك الذين رسموا الرجل مرتديا الزى الأوربى مستخدما مفتاح هاريس المعدل الذى يتألف من ٧٣ مفردة ، وحسب الثبات بطريقة تكرار المصحح الواحد ، وتكرار المصححين المختلفين ، فبلغ معامل الموضوعية فى الحالة الأولى ٩٣ر ، وفى الحالة الثانية بين ٦١ر ، ٨٧ر ، وكذلك استخدم طريقة اعادة الاختبار فبلغ معامل الثبات ٨٤ر ، ٩٨ر للاعمار المختلفة ٠ وفى حساب الصدق المرتبط بالمحكات استخدم الباحث مقياس ستانفرد - بينيه وبلغ معامل الارتباط بين ٨٠ر ، ٥٤ر ، والتحصيل الدراسى وكانت معاملات الارتباط بين ٤٢ر ، ٨١ر ٠ كما قارن الباحث بين مجموعتين متضادتين احداهما من المتخلفين عقليا والثانية من الأطفال العاديين وكانت الفروق دالة فى الاتجاه المتوقع ، كما استخدم الباحث محك تمايز العمر ، ولم يظهر البحث فروقا دالة بين الجنسين فى معظم الصفوف الدراسية ٠

ولعل من المشكلات الجوهرية التى يظهرها هذا الاختبار مسالة زى الرجل المرسوم ، وهى مشكلة حاول أن يحلها كل من مالك البدرى ومحمد متولى غنيمة عن طريق التضمين أى التساهل فى تصحيح بعض المفردات ٠ وقد قامت امال احمد مختار صادق (١٩) فى عام ١٩٧٩ بدراسة مقارنة مستقلة للمقارنة بين رسوم الرجل مرتديا الزى التقليدى السعودى والزى الأوربى ، وقد اجريت الدراسة على عينة من الأطفال قوامها ١٩٢٣ طفلا وطفلة تمتد اعمارهم بين ٨ سنوات ، ١٥ سنة ٠ وقد اختارت سن ٨ سنوات كحد ادنى لأنه العمر الذى اظهرت البيانات الامبريقية أن نوعى الزى يتمايزان فيه فى رسوم الأطفال ، وقامت الباحثة بدراسة الفروق بين النسب لكل مفردة

من مفردات المقياس فى نوعى الرسم وكانت الفروق الدالة جميعا لصالح الزى الأوربى ، وتعلق الباحثة على هذه النتيجة بقولها (١٩ : ١١٩) :

« تؤكد هذه النتيجة أن رسم الرجل بالزى الأوربى يهيىء الفرصة للطفل ان يتنبه الى تفاصيل اكثر لا ترجع الى طبيعة الزى وحده ٠٠ فكثير من المفردات التى أظهرت فروقا دالة لصالح الزى الأوربى تتضمن تفاصيل الجسم البشرى ذاته.ومن ذلك وجود الرقبة ووجود الشعر » ٠

ومن الدراسات التقنينية الهامة التى اجريت على اختبار جودانف - هاريس طبعة ١٩٦٣ تلك التى أجريت فى المملكة العربية السعودية عـام ١٩٧٧ وقام بها فريق من الباحثين باشراف فؤاد ابو حطب (٢٧) ، ولعـل اهم الاضافات فى هذا البحث أن بروتوكولات التصحيح المستخدمة (وهى من اعداد فؤاد أبو حطب) جعلت عدد المفردات ٧٧ مفردة نتيجة اضـافة مفردات جديدة أو مفردات تضمنها الاختبار الأصلى (طبعة ١٩٢٦) ٠ وقد أجرى البحث على ٢١٥ طفلا وطفلة من مرحلة ما قبل المدرسة تمتد أعمارهم بين ٣ - ٥ سنوات ، ١٩١١ طفلا وطفلة تمتد اعمارهم بين ٦ - ١٥ سـنة اختبروا جميعا من رياض الأطفال والمدارس الابتدائية والمتوسطة بالمنطقـة الغربية بالمملكة العربية السعودية ٠ وحسبت معاملات شيوع المفردات التى اظهرت أن عدد المفردات التى تعد افضل مقياس للذكاء باستخدام رسم الرجل مرتديا ثوبا (الزى الوطنى السعودى) هو ٦٦ مفردة ، أما باقى المفـردات فهى أكثر ملاءمة لرسم الرجل مرتديا الزى الأوربى ٠

وفى حساب صدق المفردات استخدمت طريقة المقارنة الطرفية وحسبت معاملات الارتباط الثنائية بين المفردة والدرجة الكلية فى الاختبار ٠ اما بالنسبة لصدق الاختبار ككل استخدم محك تمايز العمر فى تقدير صـــدق التكوين الفرضى ، والأداء فى اختبار المصفوفات المتتابعة (والذى سبق تقنينه على نفس الثقافة) لتقدير الصدق المرتبط بالمحكات ٠ اما عن ثبات الاختبار فقد حسب بطريقة اعادة الاختبار وبطريقة كيودر - ريتشاردسون ، وكانت المعاملات - كما هو متوقع - منخفضة فى الأعمار الدنيا - وخاصة اعمار ما قبل المدرسة ، اما فى الأعمار الأعلى فكانت مرتفعة نسبيا ، وكانت بالنسبة لطريقة اعادة الاختبار اقل منها بالنسبة لطريقة كيودر - ريتشاردسون ، وهى

نتيجة متوقعة فى ضوء طبيعة كل من الطريقتين فى حساب الثبات ، وتباين الخطأ المتضمن فى كل منهما ·



وتحتل مسالة الموضوعية اهمية خاصة فى اختبار رسم الرجل ، وقد حسبت بالطرق الثلاث الآتية :

١ ـ اعادة المصحح الواحد باستخدام بروتوكول التصحيح فى احدى المرتين وبدون استخدامها فى المرة الثانية ·

٢ ـ اعادة المصححين باستخدام بروتوكولات التصحيح بالنسبة لبعض المصححين وبدون استخدامها بالنسبة للبعض الآخر ·

٣ ـ اعادة المصحح الواحد باستخدام بروتوكولات التصحيح فى الحالتين ·

وكانت معاملات الموضوعية مرتفعة ، الا انها كانت منخفضة فى حالة الاختلاف فى استخدام البروتوكولات أو عدم استخدامها ·

وحسبت للاختبار معاييره فىصورة نسب ذكاء انحرافية متوسطها ١٠٠ وانحرافها المعيارى ١٥ لجميع الأعمار ·

فى عام ١٩٨٢ قام نعيم عطية (٤٧) بمحاولة بناء نسق نظرى للاختبار له طبيعة هرمية يتألف من عدة عمليات عقلية يحتاجها الأداء فى الاختبار وتشمل : ثبات الصورة الحسية ، تركيز مواقعها ، ادراك الحجم النسبى ، ادراك التناسب الداخلى، ادراك أبعاد التجسيم، الاتساق البصرى ـ الحركى، توجيه الخطوط ، الايماء والتعبير ، الاستكمال ، الرسم الجانبى · ويقترح ثلاث مراحل لاجراء الاختبار وهى :

١ ـ الرسم المواجه ، وفيها يطلب من الطفل رسم صورتين للرجل فى وضع مواجه احداهما تلقائية والثانية فى ضوء تنبيهات معينة ·

٢ ـ استقصاء الرسوم واستكمالها وتتألف من خطوتين : الاستقصاء لاستدراج التسميات والاستقصاء لاستكمال النواقص ٠

٣ ـ الرسم الجانبي ٠

ويصف الباحث بالتفصيل بروتوكولات التصحيح المقترحة لعدد من المفردات بلغ ٢٢٦ مفردة ٠ ولم يجر الباحث تقنينا شاملا للمقياس بسبب ظروف لبنان ـ وطنه ، واكتفى بدراسة امكانات التعديل الذى يقترحه للتقنين بأن طبقه على عينة صغيرة نسبيا (ن = ٣٢٩) من أطفال تمتد أعمارهم بين ٥ر٤ سنة ، ٨ سنوات ٠ وحسب الثبات بإعادة الاختبار وبلغ معامل الارتباط ٧١ر (ن = ٥٥) ، واستخدم فى حساب الصدق محك تمايز العمر (أو ما يسميه الباحث تدرج النتائج مع السن) ٠ أما المعايير فاتخذت صورة نسب الذكاء التقليدية بمتوسط ١٠٠ وانحراف معيارى ١٦ ٠

٣ ـ اختبار هايدلبرج للذكاء :

هـذا الاختبـار أعـدد فى أصـله الألمـانى كراتزيمر Kratzmeier عـام ١٩٧٧ ونقـلـه الى العـربية حسـنين الكـامل (١٦) عـام ١٩٨٥ ، ويقـوم فى جـوهره على تحليـل ما يسـمى « القـدرة التمييزية » الى خمسة مكونات هى : ادراك العلاقات ، والسلاسل المنطقية ، والانتاجية والتمييز والابتكارية ٠ وقد استبعد معد الاختبار بالعربية القسم الخامس (الابتكارية) من الصورة العربية ٠ وطبقه على عينة مكونة من ٢١٦ طفلا وطفلة من تلاميذ الصفوف الثانى والثالث والرابع من مرحلة التعليم الأساسى بمحافظة سوهاج ، وأجرى على كل اختبار فرعى (يتألف من ٧٢ مفردة) تحليلا للمفردات بحساب معاملات الصعوبة والتمييز ثم طبق الاختبار على عينة أخرى (ن = ٨٢٤) لها نفس الخصائص السابقة وحسبت لهـا المعايير فى صورة مئينيات ثم حسبت نسب الذكاء الانحـرافية ، بمتوسط ١٠٠ وانحراف معيارى ١٥ والدرجات التائية ٠

وحسب الثبات بمعامل الفا لكرونباك ٠ أما صدق التكوين الفرضى فقد حسب بطريقتين : أولاهما تمايز العمر ، وكانت الفروق فى الاتجاه المتوقع ،

والطريقة الثانية هي التحليل العاملي وتوصل الباحث الى اربعة عوامل تطابقت مع الاختبارات الفرعية فيه ۰ وحسب الصدق التلازمي باســـتخدام محك التحصيل المدرسي وكانت معظم معاملات الارتباط دالة ، وباستخدام اختبار الذكاء المصور كمحك ، وكانت معاملات الارتباط منخفضة بشكل واضح يوحي بضرورة اعادة تقنين اختبار الذكاء المصور ذاته واعادة النظر فيـه كمقياس للذكاء ۰

مراجع الباب الثالث

١ - أحمد زكى صالح : **الأسس النفسية للتعليم الثانوى** • دار النهضة العربية ، ١٩٧٢ •

٢ - أحمد زكى صالح : **علم النفس التربوى** • مكتبة النهضة المصرية ، ١٩٧٢ •

٣ - أحمد زكى صالح : **علم النفس في الادارة والصناعة** • مكتبة النهضة المصرية ، ١٩٧٢ •

٤ - احمد زكى صالح : **تعليمات اختبار الذكاء المصور** • (ب • ت)

٥ - أحمد فكرى : **الطريقة المبتكرة لقياس العقول** ، ١٩٢٠ •

٦ - اسماعيل القبانى : **مقياس ستانفرد - بينيه للذكاء** ، لجنة لتـــأليف والترجمة والنشر ، ١٩٢٨ •

٧ - اسماعيل القبانى : **قياس الذكاء** • المطبعة الأميرية ، ١٩٢٨

٨ - اسماعيل القبانى : **اختبار الذكاء الابتدائى** ، لجنة التأليف والترجمة والنشر (ط ثانية) ، ١٩٣٩ •

٩ - امال أحمد مختار صادق • الصورة السعودية لمقياس ستانفرد - بينيه (بحث غير منشور) ، ١٩٧٧ •

١٩ - ـــــــــــــ : دراسة مقارنة لاختبار رسم الرجل فى كل من الزى التقليدى والأوربى • فى بحوث فى تقنين الاختبـــارات النفسية (تحرير فؤاد أبو حطب) ، المجلد الثانى ، مكتبة الانجلو المصرية ، ١٩٧٩ ، ١٠٥ - ١٢٨ •

١٠ - أمير بقطر : **اختبارات الذكاء** ، مجلة التربية الحديثة ، ١٩٢٨ •

١١ - جابر عبد الحميد جابر ، يوسف الشيخ: **علم النفس الصـــناعى** • دار النهضة العربية ، ١٩٦٨ •

١٢ ـ جابر عبد الحميد جابر: الذكاء ومقاييسه • دار النهضة العربية ١٩٧١ •

١٣ ـ جمال محمد علي : تقنين مقياس ستانفرد ـ بينيه للذكاء على تلاميذ مرحلة التعليم الأساسي • رسالة ماجستير ، كلية التربية جامعة عين شمس ، ١٩٨٣ •

١٤ ـ حسن حسين زيتون : اختبار مراحل بياجيه للنمو العقلي • الكويت: دار القلم ، ١٩٨٦ •

١٥ ـ حسن عمر : مقياس الذكاء ، ١٩٢٨

١٦ ـ حسنين الكامل : اختبار هايدلبرج للذكاء • مكتبة النهضة المصرية ، ١٩٨٥

١١٦ ـ رجاء محمود علام ، كمال ابراهيم مرسي : مقياس ستانفرد ـ بينيه للذكاء ، الصورة الكويتية ، ادارة الخدمة النفسية لوزارة التربية ، الكويت ، ١٩٧٨ •

١٧ ـ سمية سعيد عبد القادر : دراسة تجريبية لعامل التدهور في ذكاء الذهانيين كما يبينه مقياس وكسلر ـ بلفيو • رسالة ماجستير ، كلية الآداب جامعة عين شمس ، ١٩٧٧ •

١٨ ـ سيد أحمد عثمان ، فؤاد أبو حطب : التفكير ، دراسات نفسية • الأنجلو المصرية ، (ط ١) ١٩٧٢ ، ط ٢) ، ١٩٧٧ •

٢٠ ـ السيد محمد خيرى : تعليمات اختبار الذكاء الاعدادى • دار النهضة العربية (ب • ت) •

٢١ ـ السيد محمد خيرى : تعليمات اختبار الذكاء العالى • دار النهضة العربية (ب • ت) •

٢٢ ـ صفوت فرج : القياس النفسى • دار الفكر العربى ، ١٩٨٠

٢٢ أ ـ عبد العزيز القوصى ، هدى برادة ، حامد زهران : اختبار عين شمس للذكاء الابتدائى • عالم الكتب ، ١٩٧٤

٢٣ ـ عبد الله عبد العزيز المعيلي : أساليب تقنين الاختبار وتطبيق ذلك على اختبار ستانفرد ـ بينيه ليناسب طلبة المدارس الابتدائية بالرياض رسالة ماجستير ، كلية التربية جامعة الملك سعود ، ١٤٠٢ هـ ٠

٢٤ ـ فؤاد أبو حطب : القدرات العقلية ٠ الأنجلو المصرية ، (ط ٤) ١٩٨٣

٢٥ ـ فؤاد أبو حطب (تحرير) بحوث في تقنين الاختبارات النفسية (المجلد الأول) ٠ مكتبة الأنجلو المصرية ، ١٩٧٧ ٠

٢٦ ـ ــــــــــــــ : (تحرير) بحوث في تقنين الاختبارات النفسية (المجلد الثاني) ٠ مكتبة الأنجلو المصرية ، ١٩٧٩

٢٧ ـ ــــــــــــــ وآخرون : تقنين اختبار رسم الرجل على البيئة السعودية ، في المرجع ٢٥ ، ٢٤٧ ـ ٣٤٦

٢٨ ـ ــــــــــــــ وآخرون : تقنين اختبار المصفوفات المتتابعة في المرجع ٢٥ ، ١٩١ ـ ٢٤٦

٢٩ ـ فؤاد البهي السيد : الذكاء ٠ دار الفكر العربي ، ١٩٧١

٢٩ أ ـ فتحي السيد عبد الرحيم : اختبار رسم الرجل ، تطبيقه وتقنينه على البيئة اليمنية ٠ بحث غير منشور ، ١٩٧٥

٣٠ ـ فرج عبد القادر طه : العلاقة بين الاصابات في الصناعة وكل من الصفحة النفسية للذكاء والسرعة الادراكية والسرعة الحركية ٠ المجلة الاجتماعية القومية ـ المجلد السادس ، سبتمبر ١٩٦٩ ، ٧٣ ـ ١٠٤

٣١ ـ ــــــــــــــ ، صلاح احمد مرحاب : الصورة المغربية لمقياس وكسلر ـ بلفيو لذكاء الراشدين والمراهقين المغرب : الرباط ، مطبعة الكوثر ، ١٩٧٧

٣٢ ـ فوس ، برايان (ترجمة فؤاد ابو حطب) · افاق جديدة فى علم النفس عالم الكتب ، ١٩٧٢

٣٣ ـ لويس كامل مليكة : الدلالات الاكلينيكية لمقياس وكسلر ـ بلفيو لذكاء الراشدين والمراهقين · النهضة المصرية ، ١٩٦٠

٣٤ ـ لويس كامل مليكة : المفردات فى قياس الذكاء · النهضة المصرية ، ١٩٦٠ ·

٣٥ ـ لويس كامل مليكة : نماذج التصـــحيح وجداول الدرجات الموزونة ونسب الذكاء لمقياس وكسلر بلفيو لذكاء الراشـــدين والمراهقين : النهضة المصرية ، ١٩٦٠

٣٦ ـ مائسة انور المفتى : اعداد صورة معدلة لاختبار وكسلم ـ بلفيو لذكاء الراشدين والمراهقين · رسالة ، ماجستير ، كلية الاداب جامعة عين شمس ، ١٩٧١ ·

٣٧ ـ مالك البدرى : سيكولوجية رسوم الأطفال : اختبارات رسم الانسان وتطبيقاتها على أطفال البلاد العربية · بيروت : دار الفتح للطباعة والنشر ، ١٩٦٦ ·

٣٨ أ ـ محمد بطانية : تقنين اختبار رسم الرجل على البيئة الأردنية ـ رسالة ماجستير ، الجامعة الأردنية ، فى المرجع رقم ٣٧ ·

٣٨ ـ ماهر محمود الهوارى : دراسة لصدق مقياس ستانفرد ـ بينيه للذكاء رسالة ماجستير غير منشورة· كلية الاداب جامعة عين شمس، ١٩٦٣

٣٩ ـ محمد خليفة بركات : الاختبارات والمقاييس العقلية · مكتبة مصر ، ١٩٥٤

٤٠ ـ محمد عبد السلام أحمد ، لويس كامل مليكة : مقياس ستانفرد ـ بينيه للذكاء · مكتبة النهضة المصرية ، ١٩٥٦ ·

٤١ ـ محمد عماد الدين اسماعيل ، لويس كامل مليكة : مقياس وكسلر ـ بلفيو لذكاء الراشدين والمراهقين · مكتبة النهضة المصرية ، ١٩٥٦

٤٢ ـ محمد عماد الدين اسماعيل · لويس كامل مليكة : مقياس وكسلر لذكاء الأطفال · مكتبة النهضة المصرية ، ١٩٦١ ·

٤٣ ـ محمد عماد الدين اسماعيل: بحوث فى اختبار وكسلر لذكاء الراشدين والمراهقين · مكتبة النهضة المصرية (ب · ت)

٤٤ ـ محمد متولى أحمد غنيمة : تقنين اختبار رسم الرجل بالنسبة لتلاميذ المرحلة الابتدائية فى مصر · رسالة ماجستير ، كلية التربية جامعة عين شمس ، ١٩٧٦ ·

١٤٤ ـ محمد نسيم رافت : تقنين اختبار رسم الرجل على البيئة الكويتية· الكويت ، التقرير السنوى لكتب الخدمة الاجتماعية ، ادارة التربية ، وزارة التربية بالكويت ، ١٩٦٨

٤٥ ـ مصطفى فهمى : اختبار رسم الرجل · دار مصر للطباعة ، (ب · ت)

٤٦١ ـ المقتطف : مقاييس الذكاء ، أغسطس ١٩١٧

١٤٦ ـ ميشيل اسكندر : رسوم الأطفال وعلاقتها بنموهم العقلى · رسالة ماجستير ، المعهد العالى للتربية للمعلمين (كلية التربية جامعة عين شمس) ، ١٩٣٥

٤٧ ـ نعيم عطية : ذكاء الأطفال من خلال الرسوم ، نسق جديد لاختبار رسم الرجل ، دراسة استطلاعية ، بيروت : دار الطليعة ، ١٩٨٢

٤٨ ـ هاريس ، د·ب (ترجمة محمد فرغلى فراج ، عبد الحليم محمود السيد ، صفية مجدى) ، اختبار الرسم لجودانف وهاريس (غير منشور) ، ١٩٧٦

49. Anastasi, A. Psychological Testing, Macmillan, 1961, 1958, 1976, 1982.

50. Binet, A., Simon, T. The development of intelligence in children, In : Jenkins, J.J. and Patterson, D.G. (ed.) Studies in Individual Differences. London : Methuen, 1961.

51. Bond, E.A. Tenth Grade Abilities and Achievements. Teacher College, 1940.

52. Bradway, K.P., Thompson, C.W., Gravs, R.B. Preschool IQ's after twenty-five years. J. Educ., Psychol., 1958, 49, 278-181.

53. Burt, C. Mental and Scholastic Tests. (4th ed.) London, Steples Press, 1962.

54. Carlton, J.S. and Jensen, C.M. The factorial structure of the Raven Coloured progressive Matrices. Educ. and Psychol. Mesment, 1980, 40, 111-1115.

55. Cohen, J.A. factor-analytically based rational for the Wechsler Adult Intelligence Scale. J. Consul. Psychol. 1957, 21, 451, 457.

56. Cohen, J. The factorial structure of the WAIS between early adulthood and old age. J. Consul. Psychol., 1957, 283-390.

57. Cronbach, L.J. Essentials of Psychological Testing. New York, Harper, Brothers, 1960, 1970, 1984.

58. Edwards, A.J. Individual mental testing, vol. I, Scraton : Intext Educ. Publ., 1971.

59. ———— Individual mental testing, vol. II, Scraton. Intext Educ. Publ., 1972.

60. Freeman, F.S. Theory and practice of psychological testing. (3rd. ed.), New Delhi Oxford, 1972.

61. Guertin, W.H., et al. Research with the Wechster Scales for Adults: 1960-1965 Psychol. Bull., 1966, 66, 385-409.

62. ———— et al. Research with Wechsler Scales for Adults: 1965-1970 Psychol. Rec., 1971, 21, 289-339.

63. Guilford, J.P. The Nature of Human Intelligence. McGraw-Hill, 1967.

64. Himelstein, P. Research with the Stanford-Binet, Form L.M. The first five years. Psychol. Bull., 65, 156-164, 1966.

65. Jones, L.V. A factor analysis of the Stanford-Binet at four age levels. Psychometrika, 1949, 14, 299-331.

66. Jones, L.V. Primary abilities in the Stanford Binet, age 13, J. Gent. Psychol., 1954, 74, 126-147.

67. Kennedy, W.A., et al. The ceiling of the new Stanford-Binet. J. Clin. Psychol., 1960, 17, 274-276.

68. Matarazzo, J.D. Wechsler's measurement and appraisal of adult intelligence (5th ed.) Baltimore: Williams and Wilkins, 1972.

69. McNemar, Q. The revision of the Stanford-Binet Scale : An analysis of the Standardization data. Boston : Houghton Mifflin, 1942.

70. Moriatory A.E. children's way of coping with the intelligence test. Menninger Clinic Bull., 1960, 24, 115-127.

71. Peterson, J. Early Conceptions and tests of intelligence. New York, Yonkers, Wold Book, 1926.

72. Pineau; S.R. Changes in intelligence quotient from infancy to maturity. Boston : Houghton Mifflin, 1961.

73. Sattler, J.M. Assessment of children's intelligence and special abilitie (2nd ed.), Boston : Allyn and Bacon, 1972.

74. Saunders, D.R. On the dimentionality of the WAIS battery for two groups of normal males. Psychol. Rep., 1959, 5, 529-541.

75. Saunders, D.R. A factor analysis of the information and Arithmetic items of the WAIS, **Psychol. Rep.**, 1960, 6, 367-383.

76. Saunders, D.R. A factor analysis of the Picture Completion items of the WAIS. **J. Clin. Psychol.**, 1960, 16, 146-149.

77. Saunders, D.R. Alpha frequency : A cross-validation. **J. Cli. Psychol.**, 1961, 17, 165-167.

78. Terman, L.M., Merrill, M.A. **Measurement of intelligence,** Boston, Hougton Mifflin, 1937.

79. Terman, L.M., Merrill, M.A. **Stanford-Binet Intelligence Scale. Manual for the third revision, Form L-M.** Boston : Hougton Mifflin, 1960.

80. Terman, L.M. and Merrill, M.A. Stanford-Binet Intelligence Scale : 1972 norms edition, Boston : Hougton Mifflin, 1973.

81. Thorndike, R.L., Hagen, E.P. and Sattler, J.M. Stanford-Binet Intelligence Scale, (4th ed.) Chicago : The Riverside Publishing Co., 1985.

82. Vernon, P.E. **Intelligence and attainment tests.** University of London Press, 1960.

83. Wechsler, . **The measurement and appraisal of adult intelligence** (4th ed.) Baltimore : Williams and Wilkins, 1958.

84. Wolf, T.H. **Alfred Binet. Chicago** : Univ. of Chicago Press, 1973.

البـــاب الرابع

مقـاييس الاستعدادات

الفصـــل التاسـع

بطاريات الاستعدادات المتعددة

يقصد ببطاريات الاستعدادات المتعددة multiple-aptitude batteries مجموعة المقاييس التى تهدف الى قياس مجموعة من مظاهر النشاط العقلى، كل على حدة · قياسا مستقلا ، وتعطى لكل منها درجة مستقلة · وعادة ما لا تعطى هذه البطاريات مؤشرا واحدا للبطارية ككل (مثل نسبة الذكـاء أو العمر العقلى او الدرجة المعيارية) · وفى هذا تختلف بطاريات الاستعدادات المتعددة عن اختبارات الذكاء العام التى تناولناها فى الباب السـابق ، كما تختلف عنها أيضا فى أن اختبارات الذكاء العام تعتمد فى بنائها على درجة تجانس أسئلتها أو اختباراتها الفرعية بحيث تعطى فى النهاية مقياسا للذكاء يتوافر فيه شرط الاتساق الداخلى ، حتى ولو كانت هذه الاختبـارات تعطى درجات منفصلة للاقسام اللفظية أو العملية فيها (مثل مقاييس وكسـلر) أو للاقسام اللفظية والعددية (مثل اختبارات ذكاء تلاميذ المرحلة الثانوية وطلاب الجامعات) ، بالاضافة الى الدرجة الكلية او التقدير العام للذكاء· أما فى بطاريات الاستعدادات المتعددة فان الهدف من القياس هو اعطـاء المفحوص عددا من الدرجات المنفصلة للتمييز بين القدرات العديدة التى تظهر اختلافات داخل الفرد الواحد بقدر ما تظهرها بين الافراد ·

ويمكن أن نلخص أهم العوامل التى أدت الى زيادة الاهتمام بهـذا الاتجاه « الفارق » وخاصة منذ نهاية الحرب العالمية الثانية فيما يلى :

١ ــ زيادة الاعتراف بالفروق الفردية داخل الفرد intraindividual variations فى أدائه على الأجزاء المختلفة لاختبارات الذكاء العام ــ سواء كانت أسئلة منفصلة أو اختبارات فرعية · وقد تناول أحد مؤلفى هذا الكتاب فى موضع آخر (١٩) الأدلة والشواهد التجريبية على وجود هذه الفروق داخل الفرد الواحد ·

٢ ـ زيادة الاعتقاد بأن ما يسمى اختبارات الذكاء العام أقل عمومية مما كان مفترضا فيها • وتوجد شواهد كثيرة على أن بعض هذه الاختبارات، كمقياس ستانفرد ـ بينيه ، يقيس الفهم اللفظى بالاضافة الى عدد من القدرات الأخـــرى •

٣ ـ لم تتناول اختبارات الذكاء العام بعض جوانب النشاط العقلى مثل القدرات الميكانيكية والموسيقية والكتابية ، باستثناء قليل من الاختبـارات العملية للذكاء • ولذلك سرعان ما ظهرت اختبارات «الاستعدادات الخاصة» special aptitudes ـ كما تسمى ولا زالت فى بعض المؤلفـــات المتخصصة ـ لتكمل نقائص اختبارات الذكاء العام •

٤ ـ زيادة نشاط علماء النفس فى المجـــالات التطبيقية ـ كالتوجيه والارشاد التربوى والمهنى ، والانتقاء والتصنيف فى الميـــادين الصناعية والعسكرية ـ مما تطلب توافر الاختبارات التى تقيس الاستعدادات المتعددة، بل ان ظهور بض الاختبارات المبكرة التى تقيس « الاستعدادات الخاصة » ، مثل الاستعداد الكتابى والميكانيكى ، كان انعكاسا لهذا الاهتمام •

٥ ـ التوسع فى استخدام منهج التحليل العاملى عند بناء الاختبارات العقلية • وقد أمكن لعلماء النفس بهذا المنهج ان يتوصلوا الى تحديد أدق للقدرات المختلفة التى يتضمنها ذلك المصطلح العام « الذكاء » • بالاضافة الى تصنيف هذه القدرات والوصول الى نماذج نظرية لما يسمى بالتنظيم العقلى (راجع ١٩) ولذلك ظهرت البطاريات التى تقيس « عوامل متعددة » • ويتطلب هذا بالطبع أن تتألف من مجموعة من الاختبارات ، يقيس كل منهـا عاملا من العوامل التى امكن تحديدها بالتحليل العاملى •

وسوف نتناول فى هذا الفصل طائفة ن هذه البطاريات التى تقيس عددا من الاستعدادات فى وقت واحد وتسمى لهذا السبب بطاريات الاستعدادات المتعـددة •

اختبارات القدرات العقلية الاولية :

تعد بطارية اختبارات القدرات العقلية الأولية من اهم نتائج استخدام

منهج التحليل العاملي في بناء الاختبارات • وقد اعد هذه البطارية العالم الأمريكي ثرستون Thurstone محصلة لنشاطه العلمي الواسع في ميدان تحليل القدرات العقلية • وقد نشرت في عام ١٩٤١ باسم : اختبارات شيكاغو للقدرات العقلية الأولية The Chicago Tests of Primary Mental Abilities وظلت تصدر بهذا الاسم حتى عام ١٩٤٧ • وفيها انتقى ثرستون لكل عامل الاختبارات التي تشبعت به تشبعا عاليا ، واعدت لطلاب المرحلة الثانوية والجامعات • وبلغ عدد الاختبارات ١١ اختبارا تقيس ٦ عوامل أولية هي : الفهم اللفظي (V) ، والقدرة المكانية (S) والقدرة العددية (N) والذاكرة (M) ، وطلاقة الكلمات (W) ، والاستدلال (R) • واستخدام لقياس كل قدرة اختباران ، ما عدا الذاكرة فقد استخدم في قياسها اختبار واحد ، وقد طبعت هذه الاختبارات في كراسات بحيث يمكن تطبيقها في حصتين دراسيتين (أي في حوالي ساعتين) •

ثم ادخلت تعديلات على هذه البطارية منذ عام ١٩٤٧ فاصبحت تسمى اختبارات SRA للقدرات العقلية الأولية نسبة الى الهيئة العلمية التي تصدرها •

ومن أهم هذه التعديلات اختزال طولها وخفض مستواها الى المستويات العمرية الأدنى • وأصبح لها ٣ مستويات • من ٥ ـ ٧ سنوات ، ومن ٧ ـ ١١ سنة ، ومن ١١ ـ ١٧ سنة • وقد تعرضت هذه البطارية (طبعة ١٩٤٧) للنقد في معظم مجلدات الكتاب السنوي للقياس العقلي الذي يصدره بورس، بالاضافة الى انتقادات سوبر لها (٤٥) وأهمها عدم ملاءمة بيانات التقنين ، وعدم كفاءة المعايير (وبخاصة استخدام نسب الذكاء التقليدية) ، وعدم دقة تفسير الدرجات في ضوء المحكات المهنية والتعليمية ، وعدم توافر بيانات كافية عن صدق البطارية وعدم صلاحية الطرق التي استخدمت في حساب ثبات بعض الاختبارات ، فقد استخدمت طريقة التجزئة النصفية في بعض الاختبارات دون مراعاة اعتماد درجاتها اعتمادا كبيرا على السرعة ، ونقصان معاملات ثبات بعض الدرجات العاملية (٣١ : ٢٧٤) • ومن المعروف أن افضل طرق حساب الثبات لهذه الاختبارات التي تعتمد على السرعة طريقتا اعادة الاختبار والصور المتكافئة ، فاذا اضطر الباحث الى استخدام التجزئة

النصفية فيجب أن تتم التجزئة على أساس السرعة في الأداء لا على أساس الأسئلة فى ذاتها وهذا ما اكده بحث انستازى ودوريك (٣٣) ٠

وقد تعرضت بطارية القدرات العقلية الأولية لتعديل آخر عام ١٩٦٢ للتغلب على هذه الصعوبات الفنية الا انه لا تزال اقل فى مستواها العلمى من نظائرها من البطاريات ، ناهيك عن تخلفها المعاصر عن التطورات التى طرأت على نظريات التنظيم العقلى وعدد القدرات التى يتألف منها العقل الانسانى ٠ والطبعة الراهنة لهذه الاختبارات تتألف من ٥ بطاريات على النحو الآتى :

١ ـ من مرحلة الحضانة حتى بداية الصف الثانى الابتدائى (من ٥ ـ ٨ سنوات) ٠

٢ ـ من الصف الثانى حتى الصف الرابع الابتدائى (من ٨ ـ ١٠ سنوات) ٠

٣ ـ من الصف الرابع حتى الصف السادس الابتدائى (من ١٠ ـ ١٢ سنة) ٠

٤ ـ تقابل المستويات العمرية للمرحلة الاعدادية فى مصر (من ١٢ ـ ١٥ سنة) ٠

٥ ـ تقابل المستويات العمرية للمرحلة الثانوية فى مصر (من ١٥ ـ ١٨ سنة) ٠

وتعطى كل بطارية من هذه البطاريات درجات منفصلة فى القدرات الأولية الآتية :

١ ـ الفهم اللفظى (معانى الكلمات) ٠
٢ ـ القدرة العددية ٠
٣ ـ الاستدلال ٠
٤ ـ السرعة الادراكية ٠
٥ ـ العلاقات المكانية ٠

وتتضمن بطارية الأعمار من ١٠ـ١٢ سنة اختبارات لقياس العوامل الخمس

اما بالنسبة لبطاريات المستوى الأدنى يحذف منها اختبار الاستدلال لصعوبته، وبالنسبة لبطاريات المستوى الأعلى يحذف منها اختبار السرعة الادراكية لسهولته • وتتحول درجات كل عامل ، وكذلك الدرجة الكلية في البطارية الى نسب ذكاء انحرافية بدلا من نسبة الذكاء التقليدية فى طبعة ١٩٤٧ • وتذكر كراسات التعليمات أن معلمى دور الحضانة والصفوف الأولى من المرحلة الابتدائية يهتمون بالأعمار العقلية أكثر من غيرها من المعايير ولذلك حولت جميع درجات بطارية مرحلة الحضانة حتى بداية الصف الثانى الابتدائى الى نسب ذكاء تقليدية فقط ، كما حولت درجات بطارية الصف الثانى حتى الصف الرابع الابتدائى الى نسب ذكاء تقليدية ونسب ذكاء انحرافية• وهذا المطلب واجهته ايضا كما أشرنا مقاييس وكسلر للاطفال فى المرحلة الابتدائية WISC وفى درر الحضانة WPPSI ؛ الا ان هذه المقاييس أعطت الأعمار العقلية واحتفظت بانحرافات معيارية موحدة لجميع الأعمار، وهذا ما كان يجب على اختبارات القدرات العقلية الأولية عمله بدلا من تغيير وحدات القياس فى المستويات العمرية الدنيا • وبالاضافة الى ذلك تتوافر للبطاريات الأخرى معايير مئينية وتساعية •

وقد اثمرت بطارية القدرات العقلية الأولية ثروة من المعلومات التى توافرت من البحوث الكثيرة التى أجريت عليها فى مراحلها الثلاث ، ومع ذلك فاننا نذكر التحفظات الآتية عليها :

١ ـ ان معاملات الارتباط بين الاختبارات الفرعية التى تتألف منها البطارية مرتفعة نسبيا ، فقد تراوحت هذه المعاملات لبطارية الأعمار من ٨ ـ ١٠ بين ٤١ر ، ٧٠ ر ، ولبطارية الأعمار من ٥ ـ ٨ بين ٥١ر ، ٧٣ر• وبالطبع فان هذه المعاملات المرتفعة لا تتفق مع الأساس الذى قامت عليه بطاريات الاستعدادات المتعددة ، أى الاستقلال النسبى للعوامل التى تقيسها هذه البطاريات • وقد يرجع ذلك الى عدم استقلال هذه الاستعدادات فى هذه المراحل المبكرة من العمر •

٢ ـ لا تتوافر لهذه البطارية بيانات امبريقية عن صدقها فى ضوء المحكات التعليمية والمهنية ، كما لا تتوافر بيانات عن الفروق بين الجنسين فى بعض الاختبارات التى تقيس العوامل التى اكدت الدراسات السابقة وجود

فروق بين الجنسين فيها · وبصفة عامة فان نتائج البحث فى صدق هذه البطارية فى صورها المبكرة تدل على ان اختبارى معانى الكلمات والاستدلال يرتبطان أعلى من غيرهما من الاختبارات الفرعية بمحكات النجاح المدرسى · ويعلق فريمان (٣٧ : ٤١٥) على ذلك بأن هذين الاختبارين يتشابهان مع ما تضمنته مقاييس الذكاء العام · بل ان اختبارات الذكاء العام - فى رأيه - تتضمن أسئلة تقترب من حيث التعقد والبناء من أنواع النشاط العقلى التى يتطلبها العمل المدرسى اكثر من هذين الاختبارين من بطارية القدرات العقلية الأولية ·

٣ - لم تساير بطارية القدرات العقلية الأولية التطورات التى لحقت بنظرية الذكاء فى الوقت الحاضر · ومن ذلك مثلا أن اختبار الاستدلال الذى يستخدم فيها ، وهو من نوع تكملة سلاسل الحروف ، ليس أفضل اختبارات الاستدلال العام · وكذلك فان اختبار مراجعة الجمع العددى الذى يستخدم فى هذه البطارية ليس أفضل مقاييس القدرة العددية · بالاضافة الى أن أى بطارية معاصرة للقدرات المتعددة يجب أن تتضمن عدة اختبارات لقياس الأنماط المتعددة للتفكير كما تحددت فى الدراسات العلمية الحديثة (١٩) ·

٤ - الخلط الواضح فى بنية هذه البطارية بين قدرات العمليات المعرفية (الاستدلال ، الذاكرة ، الادراك) وقدرات المحتوى (القدرة اللغوية ، القدرة المكانية ، القدرة العددية) وهو تمييز ليس له ما يبرره فى النظرية الحديثة للقدرات العقلية (١٩) ·

وقد أعد هذه البطارية باللغة العربية أحمد زكى صالح وعدلها بما يتفق مع البيئة المصرية ، وتتألف من ٤ اختبارات فرعية هى :

١ - معانى الكلمات : ويتطلب من المفحوص التعرف على الكلمة التى ترادف فى المعنى كلمة معطاة ·

٢ - الادراك المكانى : ويتطلب من المفحوص التعرف على الأشكال المنحرفة والمعكوسة أو المقلوبة ·

٣ - التفكير (أو الاستدلال) : ويتطلب من المفحوص تكملة سلاسل حـــروف ٠

٤ - العدد : ويتطلب من المفحوص مراجعة صحة عمليات جمـــع معطاة ٠

وتعطى للمفحوص درجة فى كل اختبار من هذه الاختبارات الأربع ، بالاضافة الى درجة خاصة بالنتاج العام تحسب بالمعادلة الآتية :

ق٠ع = ا ل + ½ ك + ١ ف + ١ ع ٠

حيث الرمز ق٠ع = القدرة العامة ٠

ا ل = درجة المفحوص فى اختبار معانى الكلمات ٠
½ ك = نصف درجة المفحوص فى اختبار الادراك المكانى ٠
١ ف = درجة المفحوص فى اختبار التفكير ٠
١ ع = درجة المفحوص فى اختبار العدد ٠

وتحول الدرجات الخمسة الى تخطيط نفسى (بروفيل) باستخدام معايير البطارية ٠ وقد طبقت البطارية لأكثر من ١٥ عاما على مجموعات كبيرة من الأفراد ، لا يقل عدد الأفراد الذين طبق عليهم الاختبار فى كل سن عن خمســمائة فرد من الجنسين هـم العينـة الكلية التى اســتخرجت منها معايير الاختبار ، وفى بعض الأعمار زاد عن ذلك مثل سن السابعة عشرة ، (٢ : ٩) ٠ واستخرجت المعايير لكل عمر زمنى على حدة من سن ١٣ - ١٧ سنة فى صورة مئينيات ، ونسب ذكاء تقليدية ٠

وتبين كراسة تعليمات البطارية معاملات ثبات الاختبارات الفرعيــة وهى تتراوح بين ٨١ر ، ٩٥ر ٠ الا انها لم تبين حجم العينات التى حسبت منها هذه المعاملات ولا الطريقة التى استخدمت فى ذلك ٠ كما حسبت معاملات الارتباط بين الاختبارات الفرعية ، وهى بصفة عامة معاملات منخفضة مما يدل على الاستقلال النسبى للعوامل ٠ أما عن صدق الاختبارات فتــذكر

كراسة التعليمات أنه حسبت معاملات الارتباط بين الاختبارات الفرعية وغيرها من الاختبارات بالاضافة الى التكوين العاملي للاختبارات الفرعية ، ونحن ننصح المهتم بهذه البطارية بالرجوع الى البحوث التى أشرف عليها أحمد زكى صالح حتى وفاته عام ١٩٧٤ والتى قام بها طلاب الدراسات العليا بقسم علم النفس بكلية التربية جامعة عين شمس ، للحصول على مزيد من التفاصيل · كما ننصح أى مستخدم لها فى الوقت الحاضر أن يعيد تقنين البطارية حيث أن معظم بياناتها السيكومترية تحتاج الى تحديث ·

اختبارات الاستعدادات الفارقة :

ظهرت هذه البطارية لأول مرة عام ١٩٤٧ باسم اختبارات الاستعدادات الفارقة Differential Aptitude Tests (DAT) واعدها بينيت C.K. Bennett وسيشور H.B. Seashore ووزمان A.C. Wesman ، ثم عدلت عام ١٩٦٣ وأعيد تقنينها ، وظهرت الطبعة الرابعة منها عام ١٩٦٦ ·

وتعد هذه البطارية من أكثر بطاريات الاستعدادات المتعددة طموحا ، وتتوافر عنها ـ فى الخارج ـ بيانات احصائية من خلال التقنين والتحليل فى دراسات عديدة كثيرة · ورغم أنها تصلح خاصة للمستويات العمرية من ١٤ سنة الى ١٨ سنة ، وفى أغراض التوجيه التربوى والمهنى ، الا أنها يمكن أن تستخدم مع الجماعات غير المنتقاة من الراشدين ·

وتتألف البطارية من ٨ اختبارات قننت جميعا على نفس عينة التقنين، ومعنى ذلك أن معايير الاختبارات لها نفس الدلالة النسبية · كما أمكن التحكم فىمدى العمر الزمنى والاستعداد والمستوى الدراسى والعوامل غير العقلية فى الشخصية · ومعنى ذلك أن التخطيط النفسى له معناه عند تفسير الفروق داخل الفرد · وتتوافر بطاريات منفصلة ومعايير مستقلة للذكور والاناث فى صورة مئينيات وتساعيات ·

وتقاس القدرات المتضمنة فى هذه البطارية قياسا مستقلا ، بالاضافة الى أن الاختبارات التى تتضمنها تعتمد على القوة فى الأداء وليس السرعة، فيما عدا اختبار السرعة والدقة الكتابية ·

ورغم أن مؤلفى البطارية لم يستخدموا منهج التحليل العاملى فى اعدادها ، الا انهم استعانوا فى الاختبارات التى تتألف منها بنتائج البحوث العاملية بالاضافة الى المطالب العملية ، ولذلك فانهم لم يسعوا الى اختيار الاختبارات « ذات التكوين العاملى البسيط » أو « النقية عامليا » ، وانما اختاروا الاختبارات المركبة عامليا ما دامت تفيد فى مجالات تربوية أو مهنية لها أهميتها · وتتألف البطارية من الاختبارات الآتية :

١ ـ الاستدلال اللفظى : وتتألف أسئلته من القياس التمثيلى اللفظى ويتشبع بالمفاهيم والعلاقات اللفظية المركبة ، ومع ذلك فيغلب عليه طابع الاستدلال وليس الفهم اللفظى · لأن الاختبار لا يؤكد صعوبة المفردات اللغوية، ومن أمثلة أسئلته :

··· بالنسبة الى المساء مثل الفطور بالنسبة الى ··

« ا » العشاء ـ الركن ·

« ب » الهادىء ـ الصباح ·

« حـ » الباب ـ الركن ·

« د » الجرى ـ الاستمتاع ·

« هـ » العشاء ـ الصباح ·

ويرى فريمان (٣٧ : ٤١٧) أن هذا الاختبار يقيس فى الواقع جانبا محدودا من الاستدلال اللفظى ، ولذلك يحسن أن يسمى اختبار « التماثل اللفظى » ·

٢ ـ القدرة العددية : ويشمل هذا الاختبار مدى واسعا من العمليات الحسابية ، ولا يلعب فيه حل المسائل الحسابية أى دور · وتقيس بعض الأسئلة المهارة فى اجراء العمليات الأساسية الأربعة · ويتطلب البعض الآخر فهم المفاهيم والعلاقات الكمية ·

٣ ـ الاستدلال المجرد : ويقيس القدرة على الاستدلال باستخدام المواد غير اللفظية ، ويتألف من سلاسل من الرسوم ، يتطلب كل منها فهم المبدأ الذى يؤدى الى التغير فى الأشكال التى تتألف منها السلسلة أى من نوع سلاسل الأشكال ·

٤ - العلاقات المكانية : ويقيس القدرة على التصور البصرى المكانى وذلك بعرض شكلين هندسيين من الأشكال الثنائية البعد • ويقوم المفحوص بمعالجتهما ذهنيا بحيث يؤديان الى شكل ثلاثى البعد • والغرض من الاختبار قياس القدرة على التصور البصرى للاشكال المركبة بعد تدويرها •

٥ - الاستدلال الميكانيكى : ويقيس القدرة على الفهم الميكانيكى بواسطة مواقف مصورة تمثل المبادىء الميكانيكية والعملية • ويصاحب كل صورة سؤال بسيط موجز حول المبدأ المتضمن •

٦ - السرعة والدقة الكتابية : ويقيس السرعة والـــدقة فى الاســـتجابة للاقترانات الثنائية بين الحروف والأعداد ، ويتطلب المزاوجة بين مختلف الاقترانات مؤكدا ادراك التفاصيل ومعدل الاستجابة • وتهدف الأسئلة التى يتكون منها الى ، الاقتراب من العناصر التى تتألف منها كثير من المهـــن الكتـــابية » •

٧ - الاستخدام اللغوى (القسم الأول) : ويقيس القدرة على تهجى الكلمات، وذلك عن طريق تحديد ماذا كانت الكلمات التىيتألف منها الاختبار صحيحة أم خاطئة من الناحية الاملائية •

٨ - الاستخدام اللغوى (القسم الثـــانى) : ويقيس القدرة على استخدام القواعد اللغوية ، وذلك عن طريق تحديد ما اذا كانت الجمل التى يتألف منها الاختبار صحيحة أو خاطئة من ناحية النحو وعلامات الترقيم واستخدام الكلمات •

وتتكون هذه البطارية من صورتين متكافئتين هى الصـــورة (ل) والصورة (م) • وتعطى كراسة التعليمات الخاصة بالطبعة الرابعة التى ظهرت عام ١٩٦٦ تفاصيل كافية عن الاجراءات التى استخدمت فى بنـــاء البطارية • وقد قننت فى الولايات المتحدة على ٥٠٠٠٠ تلميذا فى ١٩٥ مدرسة فى ٤٣ ولاية •

ويذكر السيد محمد خيرى (١٢ : ٣٣١) انه قام بترجمة هذه البطارية

الى اللغة العربية بالاشتراك مع لويس كامل مليكة • وتستخدم فى مصر اجزاء مختلفة كثيرة من هذه البطارية لأغراض متعددة منها اختبار الاستدلال الميكانيكى الذى قامت مراقبة الاختبارات السيكولوجية بمصلحة الكفاية الانتاجية باقتباس جزء منه وتعديله بما يتناسب مع اهداف الانتقاء لنظام التلمذة الصناعية فى حرف مختلفة ، واختبار الاستدلال اللفظى والسرعة والدقة فى الأعمال الكتابية اللذان اقتبسا وعدلا عند استخدامها فى اختبارات ديوان الموظفين الذى كان موجودا فى مصر خلال الخمسينات والستينات •

والواقع أنه لا تتوافر لنا أى بيانات تقنين عن هذه البطارية ككل فى اللغة العربية فيما عدا اختبار الاستدلال الميكانيكى ، فقد حسبت المئينيات والدرجات التائية لعينة حجمها ٢٢٣٩ تلميذا فى مراكز التدريب المهنى التابعة لمصلحة الكفاية الانتاجية والتدريب المهنى لوزارة الصناعة (١٣ : ٤٩) •

كما حسب له معامل ثبات بطريقة التجزئة النصفية (١٣ : ١٥) فبلغ ٦١ر٠ واستخدم فى حساب صدقه - كجزء من بطارية اختبارات استعدادات حرف المعادن التى سنشير اليها فيما بعد - منهج التحليل العاملى فبلغ مقدار تشبعه « العامل الميكانيكى العام » ٥٣ر٠ ، ولم يحسب معامل ارتباطه بمحك من محكات التدريب المهنى (١٣ : ٢١) •

بطارية الاستعدادات العامة :

ظهرت بطارية الاستعدادات General Aptitude Test Battery (GATB) عام ١٩٤٧ وأعدها خبراء مكتب التوظيف الأمريكى لأغراض التوجيه والارشاد فى ميدان الخدمة المدنية ، وتعتمد على افتراض ان « عددا كبيرا من الاختبارات يمكن أن يتجمع فى عدد من العوامل وأن عددا كبيرا متنوعا من المهن يمكن ان يتجمع فى فئات مهنية تبعا لدرجة التشابه بينها فى القدرات المطلوبة • وييسر هذا اختبار جميع القدرات المهنية عند الشخص فى جلسة واحدة وتفسير درجاته فى ضوء مدى واسع من المهن » (٣٧ : ٤٢٣) •

وتتألف البطارية من ١٢ اختبارا تعطى ٩ درجات عاملية ، وتتطلب

حوالى ساعتين ونصف فى الاجراء ، والعوامل التسع التى تحددت بمنهج التحليل العاملى هى :

١ - الذكاء العام : ويقاس بالدرجة المركبة من ٣ اختبارات هى : الادراك المكانى الثلاثى ، والمفردات اللغوية ، والاستدلال الحسابى ٠

٢ - الاستعداد اللفظى : ويقاس باختبار المفردات الذى يتطلب من المفحوص التعرف على معانى الكلمات (مترادفات) واضدادها ٠

٣ - الأستعداد الحسابى : ويشمل اسئلة تتطلب اجراء العمليات الحسابية وحل المسائل الحسابية من النوع المعتاد ٠

٤ - الادراك المكانى : ويقيس القدرة على التصور المكانى للاشياء ذات الأبعاد الثلاثة حين تعرض عرضا ثنائيا ٠

٥ - ادراك الشكل : ويقيس القدرة على المزاوجة بين رسوم لآلات معينة من ناحية وأشكال هندسية من ناحية اخرى ٠

٦ - الادراك الكتابى : ويتطلب المزاوجة بين الأسماء ٠

٧ - التآزر الحركى : ويقيس القدرة على التآزر ، وذلك بأن يطلب من المفحوص وضع علامات معينة بالقلم الرصاص فى سلسلة من المربعات ٠

٨ - مهارة الأصبع : ويقيس كفاءة المفحوص فى تجميع الأشياء الصغيرة وفكها ٠

٩ - المهارة اليدوية : ويقيس المهارة اليدوية باستخدام اليدين معا ، ثم اليد المفضلة ، فى وضع العصى الخشبية فى لوحة العصى ، ثم قلب اوضاع العصى ٠

وتتطلب الاختبارات الأربع التى تقيس عاملى مهارة الاصبع والمهارة

اليدوية استخدام بعض الأجهزة البسيطة ، أما الاختبارات الثماني الأخرى فهي من نوع الورقة والقلم · وتوجد صور متكافئة للاختبارات السبع الأولى التي تستخدم في قياس العوامل الست من الذكـــاء العـام حتى الادراك المـكانى ·

وتتحول الدرجات العاملية التسع في بطارية الاستعدادات العـامة الى درجات معيارية معدلة متوسطها ١٠٠ وانحرافها المعيارى ٢٠ ، واشتقت هذه المعايير الأمريكية من عينة حجمها ٤٠٠٠ مفحوصا تمثل المجتمع الاحصائى السكانى العام من حيث العمر والجنس والمستوى التعليمى والتوزيع المهنى والجغرافى ، وباختبار جماعات من الموظفين والمتقدمين للوظائف والمتدربين فى برامج للتدريب لمختلف المهن · وقد أمكن تحديد أنماط الدرجات التى تبين الاستعدادات الفارقة ، والحد الأدنى المطلوب من الدرجة المعيارية لكل مهنـة من المهن ، فمثلا وجد أن مهنة المحاسبة تتطلب درجة ذكاء مقدارها ١٠٥ ، ودرجة استعداد عددى مقدارها ١١٥ · أما مهنة السمكرة فتتطلب أن يكون الحد الأدنى من الذكاء ٨٥ ، ودرجة مقدارها ٨٠ فى الاستعداد العـــددى والاستعداد المكانى والمهارة اليدوية · وهكذا يمكن المقـــارنة بين بروفيل الدرجات المعيارية للمفحوص فى جميع المهن التى يصل الى درجـات القطع عندها أو يتفوق عليها ، ويستفاد بهذه المقارنة فى أغراض التوجيه والارشاد المهنى ·

وقد قامت وحدة بحوث الدراسات النفسية بالمركز القومى للبحـــوث الاجتماعية والجنائية (٢٨) بنقل هذه البطارية الى اللغة العربية ، وتم اختيار ٧ من اختبارات البطارية الاثنى عشر لتقنينها على عينات مصرية هى : اختبار العدد (العمليات الحسابية) ، واختبار ادراك العلاقات المكانية ثلاثية البعد ، والمفردات ، واختبار المزاوجة بين الأدوات ، واختبار الاستدلال الحسابى ، واختبار المزاوجة بين الأشكال · واختبار رسم العلامات · وطبقت على عينة من تلاميذ السنة الثالثة الاعدادية ، وتلاميذ المدارس الثانوية العامة والثانوية الفنية ، وتلاميذ مراكز التدريب المهنى التابعة لوزارة الصناعة بلغ عددها ٤٩٨ تلميذا ، واستخدمت درجات هؤلاء التلاميذ فى تحليـــل الأســـئلة التى تتألف منها الاختبارات السبع فحسبت معاملات التمييز ، ودرجات الصعوبة،

ثم طبقت البطارية فى صورتها المعدلة على عينة تقنين كبيرة نسبيا بالمقارنة ببحوث التقنين التى أجريت على الاختبارات المصرية من النصف الثـــالث الاعدادى والصف الثالث الثانوى (علمى وادبى) ، والتعليم التجـــــارى والصناعى ، والزراعى ، ومراكز التدريب المهنى التابعة لوزارة الصناعة ·

وحسبت معاملات الثبات بطريقة اعادة الاختبار فتراوحت بين ٧٩ر، ٩٣ر بالنسبة لمختلف الاختبارات الفرعية داخل نطاق أنواع التعليم المختلفة، كما استخدمت طريقة التجزئة النصفية لاختبارات القوة فتراوحت معاملات الثبات بين ٨١ر ، ٩٣ر ، وطريقة كيودر ـ ريتشادرسون لاختبارات السرعة فتراوحت معاملات ثباتها بين ٧٥ر ،٩٦٠ر · وحسبت المئينيات للعينات الفرعية التى تتألف منها عينة التقنين الكلية ·

وقد حسب صدق التكوين الفرضى عن طريق التحقق من قدرة اختبارات البطارية على التمييز بين أداء تلاميذ الفئات المختلفة من التعليم · واستخدمت لهذا الغرض مصفوفة متسلسلة من النسب الحـــــرجة لتحديد قيمة الدلالة الاحصائية للفروق بين المتوسطات · كما حسب الصدق التنبؤى على عينة من ٢٠٠ تلميذ من مراكز التدريب المهنى بأنواعها المختلفة وقورنت درجاتهم فى البرامج التدريبية التى تعرضوا لها قبل تطبيق الاختبارات عليهم وبعده · وعرضت النتائج فى صورة جداول توقع لكل اختبار على حدة ، ومعنى ذلك أن درجات هذا الاختبار يمكن أن تفسر فى ضوء كل من المعيار والمحك ·

اختبارات فلاناجان لتصنيف الاستعدادات :

تمثل هذه البطارية اتجاها مختلفـا فى بناء بطاريات الاســـتعدادات المتعددة ، وتسمى **Flanagan Aptitude Classification Tests (FACT)** .

وقد ظهرت نتيجة لبحوث فلاناجان حول اعداد اختبارات تصنيف أفراد القوات الجوية بالجيش الأمريكى أثناء الحرب العالمية الثانية ، ولذلك فانها تتجه وجهة عملية تطبيقية ، وتهتم بمجال التوجيه المهنى : وقد اعتمدت على منهج تحليل العمل لكثير من المهن الذى أدى الى تحديد ٢١ « عنصرا مهنيا

حاسما ، أو قدرة تميز بين الناجحين والفاشلين من العاملين فى كل مهنة · ويتميز أي عنصر مهنى أو قدرة بأنه على قدر من العمومية ، اى يشترك فى انماط عديدة من المهن ، وفى نفس الوقت على قدر من النوعية بمعنى انه يقيس شيئا مختلفا عن العناصر المهنية الأخرى ·

وتتألف البطارية من ٢١ اختبارا لقياس العناصر المهنية منها ١٩ اختبارا كتابيا ، واختباران عمليان هما النقش Carving والنقر Tapping · وهذه الاختبارات تمثل بطارية غير متجانسة تتضمن وظائف سلوكية مختلفة، منها الادراك الحسي والنشاط الحركي والذاكرة والنشاط اللغوى والعـــد والمفردات اللغوية العامة والاستدلال اللغوى والاستخدام اللغوى · وفيما يلى عرض موجز لهذه العناصر التسعة عشر التى تقيسها الاختبـــارات اللفظية ·

١ ـ الفحص inspection : وتقيس القدرة على تحديد الأخطـاء بسرعة ودقة فى سلسلة من رسوم الأشياء (وهو اختبار للادراك البصرى للتفاصيل) ·

ـ التجميع assembly : ويقيس القدرة على التصور البصرى لشيء معين حين توضع مجموعة من الأجزاء معا (وهـــو اختبار لادراك العلاقات المكانية الثلاثية البعد) ·

٣ ـ الحكم والفهم judgment & comprehension : ويقيس القدرة على القراءة مع الفهم واستخدام الحكم السليم فى المواقف العمليـة (اختبار لفهـم معـــانى الفقـرات) ·

٤ ـ الحذق ingenuity : ويقيس القدرة على الابتكار والاختراع فى اعداد اجراءات أو الات أو عروض ذكية (اختبار استدلالى يعتمد على مشكلة محددة) ·

٥ ـ اليقظة alertness : ويقيس القدرة على تقدير الموقف وملاحظة ما ينشأ فيه من اخطار أو ما يحتاجه من وسائل وخطط (اختبار لادراك التفاصيل فى صورة معينة وما بينها من علاقات) ·

٦ ـ الترميز coding : ويتطلب السرعة والدقة فى ترجمة معلومات مكتبية الى لغة شفرية معينة .

٧ ـ الذاكرة memory : ويقيس القدرة على استدعاء الرموز المتعلمة من اختبار الترميز السابق .

٨ ـ الدقة precision : ويتطلب السرعة والدقة فى اصدار حركات دقيقة للاصـــبع .

٩ ـ المقاييس Scales . ويتطلب السرعة والدقة فى قراءة المقاييس والرسوم البيانية .

١٠ ـ التآزر Coordination : ويقيس القدرة على التآزر بين حركات اليد والذراع .

١١ ـ الحساب arithmetic : ويتطلب الكفاءة فى العليات الحسابية الأساسية .

١٢ ـ الأنماط patterns : ويقيس القدرة على محاكاة واسترجاع الخطوط الرئيسية لنمط بسيط مرسوم .

١٣ ـ المكونات components : ويقيس القدرة على تحديد الأجزاء الهامة فى مجموعة من الرسوم والتخطيطات .

١٤ ـ الجداول tables : ويقيس القدرة على قراءة نوعين من الجداول اولهما يستخدم الأعداد ، والآخر يستخدم الكلمات والحـــروف الأبجدية .

١٥ ـ الميكانيكا mechanics : ويقيس القدرة على فهـــم المبـادىء الميكانيكيـــة .

١٦ ـ التعبير expression : ويقيس القدرة على نقل الأفكار كتابة او شفويا .

١٧ ـ الاستدلال reasoning : ويقيس القدرة على حل المشكلات والتعبير عن الحل فى صورة رياضية ٠

١٨ ـ المفردات vocabulary : ويقيس القدرة على فهم معـــانى الكلمـــات ٠

١٩ ـ التخطيط planning : ويتطلب الكفاءة فى تنظيم البيانات وذلك باعادة ترتيب خطوات معينة لحل المشكلة ٠

ورغم أن الاختبارات التى تتالف منها البطارية (وعددها ٢١ كما ذكرنا) تصحح منفصلة ، فان درجات مختلف الاختبارات تجمع بعضها الى بعض عند التنبؤ بالنجاح فى مهن معينة ، وتتفاوت هذه المهن من المســتوى المهنى الرفيع (كالطب والهندسة والتدريس) الى مستوى العامل الماهر ٠ وحيث أن هذه الاختبارات المنفصلة تمثل تنوعا واختلافا فى العمليات النفسية التى تقيسها ، بالاضافة الى انخفاض معاملات الارتباط فيما بينها ، فان جمع درجات هذه الاختبارات يصبح ممارسة تثير الكثير من التساؤل والشك ٠

وتتطلب بطاريات الاختبار اللفظية زمنا كليا طوله عشر ســاعات ونصف تقسم الى ٣ جلسات اختبارية ٠ وقد أعدت معاييرها الأمريكية على عينــة كبيرة من طلاب المرحلة الثانوية ، ومع ذلك فانها يعوزها التمثيل كما تقرر انستازى (٣١ : ٣٤٢) ٠

ويرى فلاناجان أن منهج العناصر المهنية يقع فى منزلة متوسطة بين تحليل القدرات الى عوامل أولية (أو نقية) من ناحية وطريقه عينة العمل التى تتم فيها المماثلة بالعناصر الأساسية فى الأعمال والمهن الحقيقية من ناحية اخـــرى ٠

وتؤكد نتائج البحوث التى قام بها فلاناجـــان (٢٥) أن معــاملات الارتباط بين الاختبارات الفرعية منخفضة بوجه عام حيث يبلغ وسيطها ٢٠ر فى عينة من الطلاب من سن ١٥ سنة ، و ٣١ر فى عينة من الطلاب من سن ١٨ سنة ٠ ومعنى هذا أن الصدق الداخلى للبطارية قد تحقق (أى الاستقلال النسبى للاختبارات الفرعية) ٠

وتؤكد نتائج الصدق التلازمى باستخدام محك التحصيل المدرسى عدم اللجوء الى استخدام الاختبارات الفرعية بسبب انخفاض معاملات الارتباط بصفة عامة ، كل على حدة ، فى أغراض التوجيه التعليمى · والواقع أن الدرجة المجمعة لدرجات جميع الاختبارات ادت الى معاملات ارتباط متعددة لها دلالة احصائية · الا أن الصعوبة الجوهرية فى اللجوء الى البطارية ككل تكمن فى الوقت الطويل الذى تحتاجه فى تطبيقها ·

وقد قام فلاناجان بدراسات تتبعية لأفراد عينة من طلاب الجامعة فى عدد من مجالات الدراسة ، بعد أن اختبرهم فى أثناء دراستهم الثانوية (لحساب الصدق التنبؤى) ، وبلغ وسيط معاملات الارتباط ٣٩ر٠ وكانت اقل معاملات الارتباط لطلاب الدين والخدمة الاجتماعية وأعلاها لطلاب العلوم الاجتماعية (التى شملت ٦ منها على سبيل المثال الاخصائى النفسى، ورجل القانون ، والمؤرخ) · كما حسبت معاملات الارتباط بين درجات الاختبارات الفرعية وتقديرات الرؤساء والمشرفين فى عدد من المهن، فوجد أن هذه المعاملات منخفضة بصفة عامة ، بل قد تصل فى بعض الأحيان الى الصفر أو عدم الدلالة الاحصائية ·

وقد أجريت على هذه البطارية دراستان فى مصر احدهما على طلاب كلية الهندسة (١١) والأخرى على طلاب كلية الطب (١٢٧) ·

بطارية اختبارات استعدادات حرف المعادن :

أعدت بطارية اختبارات استعدادات حرف المعادن فى مصلحة الكفاية الانتاجية والتدريب المهنى بوزارة الصناعة باشراف السيد محمد خيرى (١٣)· وتطلب اعدادها تحليل الأعمال التى يدرب عليها تلاميذ مراكز التدريب المهنى التابعة لهذه المصلحة ، وأسفرت نتائج هذا التحليل بالنسبة لجميع حرف المعادن والسيارات والتبريد على استخلاص القدرات المهنية التالية : العلاقات المكانية ، الفهم الميكانيكى ، التصور البصرى ، تذكر الأشكال ، سرعة الادراك، المعلومات أو الخبرة الميكانيكية ، تقدير الأطوال والمساحات والحجوم ، ثبات اليد ، مهارة الأصابع · وأعدت بطارية اختبارات لقياس هذه القدرات تتألف من قسمين احدهما يعتمد على ورقة وقلم والآخر من النوع العملى ، وفيما يلى وصف موجز لهذه الاختبارات ·

اختبارات الورقة والقلم : وتشمل ٨ اختبارات هى :

١ - الاستدلال اللفظى : ويقيس القدرة على الاستدلال اللفظى وادراك المفاهيم اللفظية والقدرة على التجريد ويتألف من ٣٠ سؤالا ويستغرق ١٥ دقيقة . ويتكون كل سؤال من جملة تتضمن مفهوما معينا أو علاقة معينة على المفحوص اكتشافها بالاختيار من متعدد .

٢ - اختبار الذكاء الاعدادى : للسيد محمد خيرى .

٣ - اختبار الاستدلال الميكانيكى : وهو مقتبس من بطارية اختبارات الاستدادات الفارقة التى أشرنا اليها آنفا .

٤ - اختبار المعلومات الميكانيكية . ويهدف الى معرفة مدى ألفة المفحوص بالعدد والآلات الميكانيكية ويتكون من ٥ أسئلة فى كل منها ست صور لعدد ميكانيكية بسيطة ، ويطلب من المفحوص التعرف على أسمائها بالاختيار من متعدد .

٥ - اختبار المعلومات الحسابية : ويقيس القدرة على التعامل مع الأعداد وذلك باجراء العمليات الحسابية الأساسية (الجمع والطرح والضرب والقسمة باستخدام الأعداد الصحيحة والكسور الاعتيادية) .

٦ - اختبار التصور المكانى : وهو مقتبس من بطارية الاستعدادات العامة التى اشرنا اليها آنفا .

٧ - اختبار تكميل الأشكال : وهو مقتبس من بطارية المعهد القومى البريطانى لعلم النفس الصناعى ويقيس القدرة على ادراك العلاقات الميكانيكية الثنائية البعد ، والقدرة على المزاوجة السريعة بين الأشكال واصدار الأحكام عليها بطريقة تعتمد على التفكير الاستنباطى وسرعة الادراك وتقدير المساحات ، ويتطلب فى الأسئلة الصعبة التصور البصرى .

٨ - اختبار تذكر الأشكال : من اعداد السيد محمد خيرى ، ويقيس القدرة على تذكر الأشكال الهندسية بسرعة .

الاختبارات العملية : وتشمل ٣ اختبارات هي :

٩ – اختبار التجميع الميكانيكي : وهو مقتبس من أصل ألماني ، ويقيس القدرة الميكانيكية وخاصة القدرة على ادراك العلاقات والمتعلقات الميكانيكية والتصور المكاني • ويعتمد على اختبارات مينيسيوتا للقـــدرة الميكانيكية • ويعتمد على تجميع الأشـــياء •

١٠ – اختبار مهارة الأصابع : وهو من تصميم محلي يجمع بين جهاز أو أكثر لمهارة الأصابع اختيرت من عدد من الاختبارات الموجودة بالفعل ، ويقيس القدرة على تحريك الأصابع وتناول الاشياء الدقيقة بسرعة ودقة •

١١ – اختبار ثبات اليد : ويقيس القدرة على التحـكم في حركة اليـد •

وقد قننت هذه البطارية على عينة كبيرة تتراوح أعمارها بين ١٤ ، ١٨ سنة من الحاصلين على الشهادة الاعدادية ، وكان تعدادها ، ينحصر بين ألف وثلاثة آلاف تلميذ على حسب الظروف التي طبقت فيها الاختبارات المختلفة ، (٣ : ١٠) ، وشملت تلاميذ المدارس الاعدادية والثانوية والمتقدمين لمراكز التدريب المهني وتلاميذ هذه المراكز • وحللت اسئلة الاختبارات وتحددت لها معاملات التمييز ، ورتبت حسب مستويات الصعوبة •

وقد حسب الثبات لمعظم الاختبارات بطريقة التجزئة النصفية ، امـــا بالنسبة للاختبارات التي تعتمد على السرعة وهما اختبارا مهارة الاصابع وثبات اليد فقد استخدمت معها طريقة اعادة الاختبار • مع ملاحظة أن اختبار العمليات الحسابية لم يبين معامل ثباته في التقرير المنشور • وكانت معاملات الثبات مرضية بصفة عامة •

ولحساب صدق التكوين الفرضي اســـتخدم منهج التحليل العـــاملي لاختبارات الورقة والقلم فقط واستخرج عامل عام يكشف عن مدى تشبع كل اختبار بالاستعداد الميكانيكي العام وكانت التشبعات على النحو الآتي :

الاستدلال اللفظى ٧٥ر٠ ، والذكاء الاعدادى ٥٨ر ، الاستدلال الميكانيكى ٥٢ر ، المعلومات الميكانيكية ٢٧ر٠ والعمليات الحسابية ٣٧ر ، التصـــور المكانى ٥٨ر ، تكميل الأشكال ٥١ر تذكر الأشكال ٤٧ر ٠

وتم حساب الصدق التلازمى بين الدرجة الكلية فى البطارية ومحكات التحصيل والكفاءة في مراكز التدريب المهنى فبلغ معامل الارتباط بينها وبين درجات المواد النظرية ٢١ر٠ وبين درجات المواد العملية ١٨ر وكانت معاملات الارتباط بين اختبارات الورقة والقلم والمواد النظرية ١٩ر ، والارتباط بين الاختبارات العملية والمواد العملية ٤٠ر

واستخرجت معايير الاختبارات الفرعية التى تتألف منها هذه البطارية من اداء عينات التقنين حيث استخرج لكل اختبار معاييره المئينية ودرجاته التائية ٠ ومما يجدر ذكره أن هذه البطارية لم يجر لها أى تقنين جـــديد منذ ذلك الحين ، ومع ذلك لا تزال تستخدم حتى الآن فى مصلحة الكفاية الانتاجية والتدريب المهنى ، وهو أمر يشكك فى جدواها فى اتخاذ قـرارات الانتقاء بعد مرور ما يقرب من ١٥ عاما على أخر دراسة منظمة لها ٠

بطارية الاختبارات الفارقة لمورسبى :

أعد هذه البطارية فى الأصل العـــالم الانجليـــزى ج٠و٠ مورسبى وأصدرتها هيئة البحوث التربوية فى انجلترا وويلز (NFER) وتتألف من الاختبارات الفرعية الآتية :

١ ـ اختبار المسلسلات المركبة : وفكرته اشبه باختبارات المصـــفوفات المتتابعة لرافن ٠

٢ ـ اختبارات القدرة العامة اللفظية والعددية والادراكية ٠

٣ ـ اختبارات السرعة وتتألف من ٦ اختبارات لقياس القدرة الكتابية والمرونة العقلية والطلاقة اللفظية والطلاقة الفكرية والسرعة اليدوية والمهارة اليدوية ٠

٤ ـ اختبار الأشــكال ٠

٥ ـ اختبار القدرة الميكانيكية

وقد نقل هذه الاختبارات الى اللغة العربية وقننها على البيئة المصرية مصطفى الشرقاوى (٢٨) كما قنن هذه الاختبارات ايضا على البيئــة الجزائرية والليبية ٠

وكانت العينات المصرية تتألف من طلبة مراكز التدريب المهنى بالقاهرة، اعاد عليهم تطبيق البطارية وبلغت معاملات الثبات لاختبار المسلسلات المركبة ٨٥ر ، والقدرة العامة اللفظية ٦٧ر ، والقدرة العامة العددية ٧٥ر ، والقدرة العامة الادراكية ٧١ر ٠ اما اختبارات السرعة فتراوحت معاملات ثباتها بين ٦٩ر للطلاقة الفكرية ، ٩٧ر للطلاقة اللفظية ، واختبار الأشكال ٨١ر ، واختبار القدرة الميكانيكية ٩٣ر ٠

وحسب الصدق للاختبارات البطارية باستخدام طريقة المقارنة الطرفية وتبين ان الاختبارات تميز بين التلاميذ الممتازين والضعاف فى دراســتهم بمراكز التدريب المهنى فى بعض الحرف ٠ وحسبت المعايير من أداء عينــة تتألف من طلاب المدارس الثانوية العامة بالقاهرة وتلاميذ دبلوم مراكز التدريب المهنى بالقاهرة وطلاب السنة الأولى بكلية الاداب والعلوم الاجتمــاعية بجامعة وهران بالجزائر ، وطلاب المدارس الثانوية بمكتب تعليم طرابلس بليبيا تمتد اعمارهم بين ١٧ ـ ١٨ سنة ٠ واستخدم الباحث معيار نسبة الذكاء الانحرافية فى معظم الاختبارات ٠

بطارية الاختبارات المعرفية مرجعية العوامل :

هذه البطارية من اعداد هيئة خدمات الاختبارات التربوية (ETS) فى برنســتون بالولايات المتحــدة الأمريكيــة وتســمى بالانجليــزية Kit of factor-referenced cognitive tests وظهرت منها عدة طبعات اخرها ما صدر عام ١٩٧٦ ، وقد نقل هذه البطارية الى اللغة العربية فريقان من الباحثين : الأول يتكون من سليمان الخضرى وانور الشرقاوى ونادية عبد

السلام بجامعة عين شمس) والثانى على حســـين بدارى وانور رياض عبد الرحيم (بجامعة المنيا) ولم يتعد الأمر فى حدود علم مؤلف هذا الفصل - حدود الترجمة أو التعريب ، ولا تتوافر لدينا أى معلومات سيكومترية عن البطارية أو مكوناتها ٠

تقويم بطاريات الاستعدادات المتعددة :

من الواضح أن بطاريات الاستعدادات المتعددة - التى تناولنا نماذج منها فى هذا الفصل - تختلف فى الأسس [illegible] عن اختبارات الذكاء العام٠ فاختبارات الذكاء العام تتضمن عينات من النشاط العقلى التى لها اهميتها فى المواقف التى تتلاءم معها هذه الاختبارات ، اما اختبارات الاستعدادات المتعددة فانها تسعى الى قياس العمليات العقلية المحدودة النطاق والتى تعد ضرورية فى مقررات دراسية معينة أو فى أنماط معينة من المهن ٠ وقد أدى اهتمام بعض علماء النفس « بالتكوين العاملى البسيط » الى اســـتخدام الاختبارات التى تتكون من أسئلة محدودة القيمة فى قياس النشاط العقلى المعقد والضرورى فى عمليات التعلم المدرسى وفى النشاط المهنى ٠ وللتغلب على هذه الصعوبة لجأ بعض صناع بطاريات الاستعدادات المتعددة الى جمع درجات الاختبارات الفرعية التى تتألف منها هذه البطاريات بطرق مختلفة ولمختلف الأغراض ٠ وفى هذا نوع من الاعتراف بأن العوامل النوعية او القدرات الخاصة لا تعمل منفردة ، وانما تنتظم مع غيرها من القدرات فى نمط من النشاط العقلى على درجة من التعقد النسبى ٠

وتؤكد نتائج البحوث أنه عند استخدام الاختبـــارات فى الأغراض التطبيقية - كالتوجيه والانتقاء والتصنيف فى مؤسسة معينة ولمهن خاصة - فان الاختبارات التى تعد خصيصا لذلك تتفوق على الاختبارات المقننة والمعدة للاستخدام العام ٠ ويصدق هذا أيضا على الميدان التربوى ، ووراء هذا الرأى ذلك الاهتمام المتزايد فى الوقت الحاضر بالاختبارات التحصيلية ٠ وكذلك فان الاختبارات التى تقنن على عينات متنوعة من فئات مهنية وتعليميـــة مختلفة تفيد اكثر من غيرها فى هذه الأغراض العملية ٠ وهـــذا ما حدث بالفعل فى تقنين معظم بطاريات الاستعدادات المتعددة ٠

وحينما يستخدم الأخصائى النفسى بطاريات الاستعدادات المتعددة فى

التمييز بين الأفراد، فى مدى صلاحيتهم لانواع التعليم والمهن المختلفة فعليه أن يتيقن من الصدق الفارق differential validity للبطاريات قبل استخدامها فقد يجد أن بعض الاختبارات الفرعية اقدر على التنبؤ بنتائج معينة من غيرها وهنا يقع صانعو هذه البطاريات فى خطا فادح حين يؤكدون فى تقنينها الفروق بين متوسطات درجات الاختبارات الفرعية التى تتألف منها بالنسبة للجماعات المهنية والتعليمية المختلفة ، ويهملون ما يمكن أن يوجد بين درجات هذه الجماعات من تداخل ، فبين طلاب الطب والهندسة مثلا على الرغم من اختلافهما كثير من العناصر المشتركة ، ومن هنا يتوافر لهم انماط متشابهة من اختبارات العوامل المتعددة ، الا أن بينهم ايضا فروق فى السمات العقلية وغير العقلية التى لا تقيسها هذه البطاريات ·

الفصـل العاشر

اختبارات القدرات المتخصصة

اشرنا فيما سبق الى أن اختبارات الذكاء العام لا تشمل فيما تقيسه بعض « القدرات العقلية » ، وقد ظهر هذا النقص حتى قبل الاهتمام ببطاريات الاستعدادات المتعددة ، وتمثل ذلك في ظهور الاختبارات التي تقيس ما كان يسمى « الاستعدادات الخاصة » ، وأشهرها واكثرها تبكيرا اختبـــارات « الاستعداد الميكانيكي » • ثم تزايدت الحاجة الى هذه الاختبارات التي تقيس ما يمكن أن نسـميه « القدرات المتخصصة Specialized abilities بزيادة الاهتمام في علم النفس التطبيقي بمشكلات التوجيه والانتقاء والتوزيع والارشاد النفسي • فتكاثر عددها وشملت مجالات اكثر تنوعا ومن ذلك ميادين النشاط الموسيقي والفني والمهني والكتابي ، بالاضافة الى النشاط الميكانيكي الذي اشرنا اليه •

وقبل أن نتناول نماذج من اختبارات القدرات المتخصصة بشيء من التفصيل لابد أن نوضح مكانة هذه الفئة من الاختبارات في النظرية الحديثة للقدرات العقلية • وبايجاز نقول أن هذه الاختبارات ظهرت في وقت كان الاهتمام كله مركزا على ما يسمى « اختبارات الذكاء العام » ، بهدف أن تكون مكملة لمفهوم « نسبة الذكاء » في وصف النشاط العقلي للمفحوص • الا أنه مع التقدم في مناهج البحث وفي نظرية القدرات العقلية • وخاصة بعـــد شيوع استخدام منهج التحليل العاملي والنماذج المؤسسة عليـه (١٩) • أصبح من الشائع اعتبار مفهوم « الذكاء » ذاته يتألف من عـــدد كبير من « القـــدرات العقليـــة » يصـــل في نمـــوذج جيلفـورد (١٩ ، ٣٨) الى اكثر من ١٢٠ قدرة ويزيد عن ذلك كثيرا في النمـــاذج المعاصرة لتجهيز المعلومات (١٩) وزاد نتيجة لذلك الاهتمام ببطاريات الاستعدادات المتعددة • بل ان بعض الاختبارات التقليدية « للاستعدادات الخاصة » أصـــبحت من مكونات بعض بطاريات الاستعدادات كما اشرنا في الفصل السابق •

والسؤال الذى تطرحه انستازى (٣١) هو : ما هــو موضــع هــذه الاختبارات فى الوقت الحاضر ؟ الواقع ان لهذه الاختبارات وظيفتين هامتين فى التقويم النفسى المعاصر هما :

١ ـ توجد بعض مجالات السلوك الانسانى لا تتوافر لها مقاييس جيدة فى بطاريات الاستعدادات المتعددة ومن ذلك الادراك البصرى والســمعى والمهارة الحركية والمواهب الموسيقية والفنية والسلوك الابتكارى · وقــد يكون السبب فى ذلك أن المواقف التى تتطلب استخدام هـــذه الاختبارات على درجة كبيرة من النوعية بحيث لا تبرر ادماجها فى بطاريات الاستعدادات المتعددة · ومعنى ذلك أن هذه الاختبارات ستظل فئة متميزة من الاختبارات العقلية حتى تتوافر بطاريات « شاملة » لقدرات الانسان · وهذا ما نتنبأ بحدوثه فى المستقبل القريب لعلم القياس النفسى ·

٢ ـ تستخدم اختبارات القدرات المتخصصــة حتى فى الأحوال التى يوجد لها ما يقابلها فى بطاريات الاستعدادات المتعددة ، ومن ذلك اختبارات القدرات الميكانيكية أو الكتابية · والســـبب فى ذلك أن بعض اختبارات القدرات المتخصصة تتوافر لها بيانات تقنين وافية ، بالاضافة الى المرونة فى استخدام الاختبارات ، وشمول الميدان الذى تقيسه · فبدلا من استخدام اختبار واحد للقدرة الميكانيكية فى احدى البطاريات المتعددة الاختبارات ، يمكن للباحث أن يطبق بطارية لاختبارات القدرة الميكانيكية اذا تطلب الأمر قياس مختلف جوانب هذه القدرة فى المفحوصين ·

ويتوافر فى الوقت الحاضر عدد هائل من الاختبـــارات التى تقيس « القدرات المتخصصة » ابتداء من المستوى الحاسى (السمع والبصر) حتى مستوى التفكير الابتكارى والنشاط المهنى · ولا يتسع المقام لتناول هـــذه الاختبارات بأى نوع من العرض الشامل ولذلك سوف نقتصر على طائفة من هذه القدرات وما يقيسها من اختبارات ·

اختبار القدرات الحركية :

ظهرت الاختبارات التى تقيس خصائص الاستجابات الحركية كالسرعة

والتآزر منذ فترة طويلة ، ويهتم معظمها بالمهارة اليدوية Manual dexterity وقليل منها يهتم بحركات الساق والقدم والتى تتطلبها بعض المهن والأعمال. ويقيس بعضها مزيجا من الاستعدادات الحركية والادراكية والمكانية والميكانيكية . وفى أغلب الأحوال فان اختبارات القدرات الحركية تتطلب أجهزة فى قياسها رغم ظهور بعض اختبارات الورقة والقلم فى أغراض التقويم النفسى الجماعى . وقد أشرنا فى الفصل الخامس الى بطاريتى فلاناجان والاستعدادات العامة اللتين تضمنتا بعض الاختبارات الحركية الجماعية الا أن الدراسات التى قام بها فلشمان (٣٦) وملتون (٤١) تؤكد عدم وجود علاقة بين اختبارات الورقة والقلم واختبارات الأجهزة التى تقيس نفس القدرات الحركية ، وفيما يلى أمثلة للاجهزة التى تستخدم فى قياس النشاط الحركى :

١ ـ مقياس القوة العضلية (الدينامومتر) Dynamometer وتستخدم فى قياس قوة قبضة اليد .

٢ ـ جهاز قياس زمن الرجع .

٣ ـ اختبار Crawford Small Parts Dexterity ويقيس المهارة اليدوية البسيطة .

٤ ـ اختبار Purdue Pegboard ويقيس النشاط الحركى الدقيق والنشاط الحركى الغليظ .

٥ ـ اختبار Stromberg Dexterity ويقيس حركات اليد والذراع .

اختبارات القدرات الميكانيكية :

ان ما يسمى القدرة الميكانيكية ليس قدرة موحدة لا تقبل التحليل ، وانما هى فى الواقع مركب من القدرات الحسية والادراكية والحركية ، بالاضافة الى قدرات ادراك العلاقات المكانية ، واكتساب المعلومات الميكانيكية ، وفهم العلاقات الميكانيكية .

ولذلك فان الاختبارات التى تقيس القدرة الميكانيكية انما تهتم بمستوى من الأداء اعلى مما تتضمنه الاختبارات الحسية والادراكية (التمييز السمعى والبصرى والحركي والتآزر العضلى والمهارة اليدوية) .

ويعد اختبار التجميع assembly الذى اعده ستنكويست Stenquist عام ١٩٢٣ اول اختبار ظهر فى هذا الميدان يقيس قدرة المفحوص على تركيب أجزاء الأجهزة الميكانيكية مثل جرس الدراجة ، أو قفل ، أو مصيدة فئران . ويتألف من ٣ سلاسل تستخدم مع مستوى عمرى يمتد بين الطفولة والرشد .

وقد عدل هذا الاختبار فى جامعة مينيسوتا عام ١٩٣٠ واصبح يطلق عليه اختبــــار مينيســـوتا للتجميع الميكانيكى Minnesota Mechanical Assembly Test واضيفت اليه أجهزة جديدة . ويصحح هذا الاختبار كسابقه فى ضوء معدل الاستجابة ودقتها . وقد اكدت الدراسات التى أجريت عليه انه علىدرجة كافية من التنبؤ بنجاح طلاب التعليم الفنى المبتدئين فى اعمال الورشة ، كما ارتبط بالنجاح فى بعض المهن الميكانيكية .

وفى عام ١٩٣٠ ظهر اختبار آخر فى جامعة مينيسوتا هو اختبـــار مينيسوتا للعـــــلاقات المكانية Minnesota Spatial Relations Test ويتألف من سلسلة من ٤ لوحات خشبية يتكون كل منها من ٥٧ قطعة خشبية مختلفة الأشكال ، وبعضها غير مألوف ويطلب من المفحوص أن يضع هذه القطع فى اماكنها فى فتحات اللوحة الخشبية ، وفى راى كثير من النقاد أن هذا الاختبار يقيس السرعة والدقة فى الاستجابة لتفاصيل العلاقات المكانية، وقدرة المفحوص على التعامل مع الأشياء والمواد المحسوسة . وبعبارة أخرى فانه لا يقيس القدرة على حل المشكلات ذات الطبيعة الميكانيكية . كما لا يصلح لقياس القدرة على معالجة شياء الصغيرة بدقة .

وفى عام ١٩٤٨ ظهر ختبار مينيسوتا للوحات الأشكال الورقية Revised Minnesota Paper ormboard

يعرض على المفحوص على ورقة مطبوعة نفس المشـــكلات التى كانت تعرض عليه فى لوحات الأشكال الحقيقية . ففى كل سؤال من اسئلة هذا الاختبار يعرض على المفحوص جزءان أو اكثر من شكل هنـــــدسى بحيث

لو جمعت معا بطريقة صحيحة فانها تكون الشكل الكلى الكامل • وعلى المفحوص تحديد الشكل الذى يمكن أن يجمع مع الأجزاء الأخرى من بين ٥ بدائل اختيارية • وهكذا يقيس هذا الاختبار القدرة المكانية كما تحددت فى البحوث العاملية للقدرات العقلية • وتؤكد نتائج البحوث التى اجريت على هذا الاختبار انه يرتبط ارتباطا دالا بنوع الأداء الميكانيكى والنجاح فى الرسم الهندسى والهندسة الوصفية •واستخدمت فى حساب صدقة محكات مختلفة مثل درجات مقررات الهندسة والمواد الفنية والورش الميكانيكية وتقديرات المشرفين ومقاييس الكفاءة الانتاجية • وبوجه عام يمكن القول ان طلاب الهندسة والعاملين بالمهن الميكانيكية يحصلون على درجات فى هذا الاختبار اعلى من غيرهم ، وقام بتقنينه على البيئة المصرية فؤاد ابو حطب وامال صادق (٢١) •

وتوجد فئة أخرى من اختبارات القدرة الميكانيكية تقيس المعلومات الميكانيكية والاستدلال الميكانيكى والفهم الميكانيكى • وهذه جميعا تتطلب من المفحوص الفة بالآلات الشائعة وبالعلاقات الميكانيكية ، الا انها لا تفترض فى المفحوص توافر معلومات فنية أكثر مما يمكن أن توفره الحياة اليومية فى المجتمعات الصناعية • ومن أشهر هذه الاختبارات تلك التى اعدها بينت وزملاؤه (٢٢) والتى تسمى اختبارات الفهم الميكانيكى Tests of Mechnical Comprehension

وتعرض على المفحوص مشكلات ميكانيكية مصورة وعدة حلول لكل منها ، وعليه أن يختار الحل الصحيح • ولهذه الاختبارات ثلاثة مستويات من الصعوبة ، وتهدف الى قياس القدرة على فهم المبادىء الفيزيائية والميكانيكية فى مواقف بسيطة نسبيا •

ومن الاختبارات الأحدث نسبيا فى قياس القدرة الميكانيكية اختبار الدافعية الميكانيكية المسمى Mellanbruch Mechanical Motivation Tests وهو اختبار مصور يتطلب من المفحوص التعرف على اشياء معينة شائعة فى المجتمعات الصناعية المتقدمة • والاختبار الثانى هو SRA Mechanical Aptitudes Test ويتألف من ٣ اختبارات فرعية تقيس :

(أ) المعلومات الميكانيكية ٠

(ب) ادراك الاشكال والتصور البصرى المكانى ٠

(حـ) حل المشكلات الحسابية الخاصة بالأعمال الميكانيكية (فى صورة جداول ورسوم بيانية) ٠

ويجب أن نؤكد أن هذه الاختبارات التى تقيس الفهم الميكانيكى تتطلب تعديلا شاملا عند محاولة استخدامها فى البيئة المصرية ، وذلك لأنها مشبعة تشبعا كبيرا بالعناصر الثقافية السائدة فى المجتمعات التى صنعت فيها ٠

اختبارات القدرة الكتابية :

القدرة الكتابية Clerical ability - كالقدرة الميكانيكية - ليست قدرة بسيطة ، وانما هى على درجة من التعقد والتركيب ٠ ورغم أن بعض بطاريات الاستعدادات المتعددة التى تناولناها فى الفصل الخامس - وخاصة بطارية الاستعدادات الفارقة وبطارية الاستعدادات العامة - تضمنت اختبارات فرعية لقياس القدرة الكتابية فان تراث التقويم النفسى يتضمن طائفة من الاختبارات التى تقيس هذه القدرة على وجه الخصوص ، بل توجد بطاريات لقياس القدرة الكتابية ٠ وعموما نستطيع القول أن اختبارات هذه القدرة تجمعها خصائص مشتركة يوضحها الجدول الآتى الذى يلخص محتوى ٦ بطاريات للقدرة الكتابية واختباراتها الفرعية ٠ومنه يتضح أن البطاريات الشاملة تسعى الى قياس عدد من العمليات النوعية ٠ بينما تسعى البطاريات الأقل شمولا الى قياس السرعة والدقة الادراكيتين والتى تعد أحد جوانب القدرة الكتابية وليست كلها ٠ وقد يكون من العوامل الكامنة وراء قصر القدرة الكتابية على السرعة الادراكية ما تؤكده نتائج البحوث من أهمية القدرات الادراكية بوجه عام فيها ٠ الا أن « القدرة الكتابية » تشمل أنواعا متنوعة من الأعمال تختلف فى مدى القدرة ومستواها ٠ ولذلك فان مقاييسها تضمن عمليات عقلية أخرى أساسية ، بالاضافة الى بعض القدرات الحسية والمهارات اليدوية عند الضرورة ٠ ويلخص الجدول رقم (١٥) الاختبارات الفرعية التى تتضمنها ٦ بطاريات للقدرة الكتابية ٠

جدول رقم (١٥) يلخص الاختبارات الفرعية التى تتضمنها ٦ بطاريات للقدرة الكتابية ·

البطارية	الاختبارات الفرعية
بطارية ديترويت Detroit	الخط : السرعة والدقة المراجعة : السرعة والدقة الحساب البسيط السرعة والدقة فى الأداء الحركى معرفة المصطلحات التجارية البسيطة ترتيب الصور التصنيف : السرعة والدقة الترتيب الأبجدى
بطارية الاستعدادات الكتابية العامة General Clerical	المزاوجة : اكتشاف الأخطاء فى الأسماء والأعداد الترتيب الأبجدى الحساب : الجمع البسيط الحساب : تحديد اخطاء الجمع المسائل الحسابية البسيطة التهجى فهم القراءة معرفة معانى الكلمات الاستخدام اللغوى : القواعد النحوية

تابع الجدول رقم (١٥)

جدول يلخص الاختبارات الفرعية التى تتضمنها ٦ بطاريات للقدرة الكتابية ٠

البطارية	الاختبارات الفرعية
بطارية مينيسوتا	مقارنة الأعداد مقارنة الأسماء
بطارية بردو Purdue	التهجى الحساب المراجعة : السرعة معرفة معانى الكلمات النقل والنسخ : الدقة الاستدلال
بطارية التوظيف Short Employment	العمليات العددية معانى الكلمات التصنيف
بطارية تورس Turse	المهارات اللفظية المهارات العددية التعليمات المكتوبة المراجعة : السرعة التصنيف الترتيب الأبجدى

وتوجد فى المكتبة العربية بطاريتان لقياس القدرة الكتابية هما :

(١) اختبارات القدرة الكتابية وهى من اعداد محمد عبد السلام احمد (٢٤) عام ١٩٦٠ وتتألف من اختبارين هما : اختبار تصنيف الأعداد ، واختبار تصنيف الأسماء ، اعتمدا على تحليل للاعمال الكتابية قام به الدكتور محمد عبد السلام أحمد بمعاونة ديوان الموظفين عام ١٩٥٣ ٠ وقنن الاختبار الأول على عينة حجمها ٨٢٢ مفحوصا من الجنسين من الحاصلين على شهادة الثانوية العامة او دبلوم التجارة الثانوية تتراوح اعمارهم بين ١٨ ، ٢٨ سنة ، وقنن الاختبار الثانى على عينة عددها ٨٨٠ فردا من الجنسين تتراوح اعمارهم ايضا بين ١٨ ، ٢٨ سنة من حملة الثانوية العامة المتقدمين لشغل الوظائف الكتابية عن طريق اعلانات ديوان الموظفين في ذلك الوقت ٠ وأعدت للاختبارين معايير مئينية وتائية وحسب لهما معامل الثبات بطريقة التجزئة النصفية فبلغ ٩٧ر٠ (ن = ١٠٠) والواقع ان استخدام طريقة التجزئة النصفية فى حساب اختبارات القدرة الكتابية - التى تعتمد اعتمادا كليا على السرعة - يثير كثيرا من المشكلات اشرنا اليها في الفصول السابقة الا اذا كان المؤلف لجأ الى تجزئة زمن الاختبار ، وهذا ما لم توضحه كراسة تعليمات هذه الاختبارات (١٢ : ٢١) ٠ أما عن صدق الاختبارين فقد حسب مؤلفهما معاملات الارتباط بينهما وبين اختبارات المهن الكتابيـــة والذكاء الاجتماعى (ن = ٤٧ من طالبات معهد السكرتارية بالعباسية) فلوحظ أن الاختبارين يشتركان مع الاختبارات التى تقيس الجوانب الآتية : السرعة ، والدقة ، سهولة استخدام اللغة ، الألفة بالأعداد ، الذاكرة المباشرة ، المهارة اليـــدوية ٠

(٢) اختبارات المهن الكتابية : من اعداد محمد عماد الدين اسماعيل وسيد عبد الحميد مرسى (٢٦) وتقيس مكونات ثلاث هى : القدرة العددية ، السرعة والدقة ، الاستدلال اللغوى ٠ وقد حسب ثبات الاختبارات الثلاث (ن = ١٠٠ من طالبات معهد السكرتارية) بطريقة التجزئة النصفية فبلغت ٨٦ ، ٨٤ر ، ٨٥ر على التوالى ٠ ومرة أخرى تثير مسألة استخدام الطريقة الملائمة لحساب ثبات الاختبارات التى تعتمد على السرعة ٠ واستخدم فى حساب صدق الاختبارات تقديرات مادة المحاســـبة والترجمة وتقديرات المدرسين والمشرفين وتراوحت معاملات الارتباط بين ٤٢ر ، ٥٤ر٠ وحسبت المعايير المئينية للاختبارات ٠

اختبارات القدرات الموسيقية :

يوجد فى الوقت الحاضر عدد من الاختبـــارات التى تصلح لقياس القدرات الموسيقية ، قد يكون اقدمها ظهورا او اكثرها اهميـــة اختبارات سيشور التى تسمى : Seashore Measures of Musical Talent

وقد ظهرت فى صورتها الأصـــلية عام ١٩١٩ ، واستمرت شـــائعة الاستخدام فى الولايات المتحدة دون تعديل حتى عام ١٩٣٩ حيث عدلت فى صورتها التى تستخدم حتى الآن • وقد كانت هذه الاختبارات نتيجة للبحوث الواسعة النطاق فى ميدان سيكولوجية الموسيقى والتى أجريت بجامعة ايوا تحت اشراف كارل سيشور • وتتألف هذه البطارية فى صورتها الحديثة من ٦ اختبارات هى • تمييز الأصوات pitch ، شدة الصوت loudness تذكر الايقاعات rhythm ، الزمن time ، نوعية الصـــوت timbre ، تذكر الألحان tonal memory .

ويتألف اختبار تمييز الأصوات من ٥٠ زوجـــا من النغمات ، وعلى المفحوص أن يحدد بالنسبة لكل زوج ما اذا كانت النغمـــة الثانية احد أم اغلظ من النغمة الأولى وفى اختبار شدة الصوت يطلب منه أن يحدد بالنسبة لكل زوج من ٥٠ زوجا من النغمات ما اذا كانت النغمة الثانية اشد ام اضعف من الأولى • ويطلب منه فى اختبار تذكر الايقاعات ان يحدد ما اذا كان النمطان الايقاعيان اللذان يتألف منهما كل سؤال (من ٣٠ سؤالا) متشابهين او مختلفين : وفى اختبار الزمن يحدد المفحوص ما اذا كانت النغمة الثانية اطول أم اقصر من النغمة الأولى (٥٠ سؤالا) • وفى اختبار نوعية الصوت يقوم المفحوص بالحكم على النغمة الثانية بانها متشابهة أو مختلفة عن النغمة الأولى من حيث نوعية الصوت (البناء الهارمونى) ، ويتكون هذا الاختبار من ٥٠ سؤالا • اما الاختبار الأخير ، وهو تذكر الألحان ، فيتكون من ٣٠ زوجا من النماذج اللحنية ، وفى كل زوج توجد نغمة واحدة مختلفة فى النموذج الثانى ، وعلى المفحوص ان يحدد اى النغمات اختلفت ويحدد ترتيبها الرقمى •

وتصلح هذه الاختبارات للتطبيق على المفحوصين ابتداء من الصف

الرابع الابتدائى حتى أعلى المستويات التى تتطلب اعدادا موسيقيا رفيعا للمراهقين والراشدين ٠

وقد ظهرت هذه الاختبارات فى أصلها الأجنبى مسجلة على أسطوانات، وقد أعدتها باللغة العربية آمال أحمد مختار صادق (٥) مسجلة على شرائط ، وأعدت تعليمات تطبيقها ٠ كما صممت كراسة اجابة مختلفة عن الطبعة الأجنبية ٠ وحسبت معاملات الثبات بطريقة التجزئة النصفية وتصحيح سبيرمان براون على عينة من ١٦٣ طالبا وطالبة من حملة الثانوية العامة المتقدمين للالتحاق بالمعهد العالى للتربية الموسيقية وتراوحت معاملات الثبات بين ٥٨ر (اختبار الزمن) ، ٨٨ر (اختبار تذكر الألحان) ٠ أما عن صدق هذه الاختبارات فقد حسبت الباحثة معاملات الارتباط بين الاختبارات الفرعية التى تتألف منها البطارية لعينات من طلاب دور المعلمين والمعلمات (ن = ١٧١) وطلاب المعاهد الموسيقية العليا (ن = ٦٠) ، وتلاميذ المدارس الاعدادية العامة (ن = ١١١) ، وتلاميذ المعهد القومى للموسيقى (ن = ٧٤) ٠ كما حسبت معاملات الارتباط بين اختبارات سيشور واختبارات القدرات الموسيقية للاطفال (بنتلى) لعينة تلاميذ المرحلة الاعدادية العامة ٠ وتلاميذ المعهد القومى للموسيقا ٠ وكذلك معاملات ارتباطها باختبارات ونج للذكاء الموسيقى، واريجون للتذوق الموسيقى، وفارنم للتدوين الموسيقى وكاتل للذكاء العام ٠ واختبار الذكاء المصور للدكتور أحمد زكى صالح ، وتتوافر فى كراسة التعليمات بيانات كافية عن المعالم السيكومترية لهذه البطارية ٠ أما عن معايير هذه البطارية فقد حسبت المئينيات الخاصة بكل اختبار فرعى ولكل عينة فرعية من عينة التقنين (٥) كما تأكد الصدق التنبؤى لبطارية سيشور من خلال دراسة تتبعية قامت بها آمال صادق ، واكرام مطر (٧) استخدم فيها اسلوب تحليل الانحدار المتعدد وتأكد فيها ارتفاع القيمة التنبؤية لاختبارات سيشور فى مجال القبول للتعليم الموسيقى فى مصر بمقارنتها بالأساليب التقليدية التى تستخدم حتى وقتنا الحاضر فى انتقاء الطلاب الجدد فى كلية التربية الموسيقية جامعة حلوان ٠

وقد أعدت نفس الباحثة (٤) صورة مختصرة لبطارية سيشور ووصفت معالمها السيكومترية بوضوح ٠

ويوجد اختبار آخر للقدرة الموسيقية ظهر عام ١٩٦٣ أعده فى انجلترا

جرلد بنتلى Bentley خصيصا لاطفال المرحلة الابتدائية ، (من ٧ ـ ١٤ سنة) ويقيس ٤ قدرات هى : تمييز الأصوات ، تذكر النغمات ، تحليل التآلفات ، تذكر الايقاعات • ويمكن تطبيق هذه الاختبارات على مستويات عمرية أعلى مع مراعاة معايير هذه الأعمار عند تفسير الدرجات •

ويتكون اختبار تمييز الأصوات من ٢٠ زوجا من الأصوات « النقية » يستمع اليها المفحوص ثم يطلب منه أن يحدد ما اذا كان الصوت الثانى من كل زوج متشابها مع الصوت الأول أو مختلفا عنه • وفى حالة الاختلاف على المفحوص أن يحدد ما اذا كان أعلى (أى أحد) أو أسفل (أى أغلظ) من الصوت الأول • أما اختبار تذكر النغمات فيتكون من ١٠ اسئلة يتطلب كل منها مقارنة ثنائية بين جملتين لحنيتين قصيرتين ، ويتكون كل منها من ٥ نغمات ، ويطلب من المفحوص أن يحدد ما اذا كانت الجملة الثانية متشابهة مع الأولى أو مختلفة عنها واذا حدث اختلاف عليه أن يحدد ترتيب النغمة المختلفة • وفى اختبار تحليل التآلفات يطلب من المفحوص تحديد عـــدد النغمات التى يتكون منها كل تآلف (يتكون الاختبار من ٢٠ تآلفا) • أما اختبار تذكر الايقاعات فيتكون من ١٠ مفردات يتطلب كل منها مقارنة ثنائية بين جملتين ايقاعيتين ، وتتكون كل واحدة من ٤ وحـــدات ايقاعية أو ما يعادلها من النماذج الايقاعية المختلفة ، ويطلب من المفحوص تحديد ما اذا كانت الجملة الثانية فى كل سؤال متشابهة مع الجملة الأولى أو مختلفة عنها • وفى حالة الاختلاف يكون عليه تحديد رقم الوحدة المختلفة •

وقد أعدت هذه الاختبارات باللغة العربية ايضا آمال أحمد مختار صادق عام ١٩٧٢ (٦) • وظهرت اختبارات القدرات الموسيقية للاطفال مسجلة على شرائط (بدلا من الاسطوانات فى الأصل الأجنبى) واعدت لها كراسة اجابة خاصة • كما فصلت تعليمات تطبيقها • وقد استخدام المؤلف الأصلى فى حساب ثبات هذه الاختبارات طريقة اعادة الاختبار ، وبلغ معامل الارتباط ٨٤ر • واستخدمت الباحثة المصرية هذه الاختبارات فى دراستين عامليتين للقدرة الموسيقية (٦) احداهما على عينة من تلاميــذ المرحلة الاعدادية (ن = ١١١) ، والاخرى على عينة تناظرها فى العمر الزمنى من تلاميذ المعهد القومى للموسيقى (الكونسيرفتوار) عددها ٧٤ تلميذا • وتراوحت اشــتراكيات الاختبـــارات الفرعية بين ٤٢ر (لاختبـــار تذكر

الايقاعات لعينة تلاميذ المرحلة الاعدادية العامة) ، ٦٨ر (لاختبار تمييز الأصوات لنفس العينة) ٠ اما اشتراكيات الدرجة الكلية للبطارية فبلغت ٨٨ر فى الحالتين ٠

ولحساب صدق الاختبارات حسبت معاملات الارتباط بين الاختبارات الفرعية ، ومعاملات ارتباطها باختبارات سيشور التى أشرنا اليها ، وباختبارات كاتل للذكاء العام ٠ وكذلك حسبت المعايير المئينية لعينتين من التلاميذ : احداهما مجموعة ٩ ـ ١٢ سنة (ن = ١٦٢) ، والأخرى مجموعة ١٢ ـ ١٥ سنة (ن = ٣٦٣) ٠

ويوجد اختبار آخر ظهر فى انجلترا عام ١٩٤١ ثم ظهرت طبعته الجديدة المعدلة عام ١٩٥٨ ثم عام ١٩٦٢ وهو اختبار ونج للذكاء الموسيقى٠

Wing Standardized Tests of Musical Intelligence

ويصلح للتطبيق ابتداء من سن ٨ سنوات ، ويختلف عن اختبار سيشور فى أنه لا يؤكد الجانب « الذرى » التحليلى كما يتمثل فى عناصر الاحساس البسيط السائد فى اختبارات سيشور ، ولذلك فانه يعتمد على محتوى موسيقى واقعى حيث تستخدم موسيقى البيانو فى كل اختبار من الاختبارات السبع التى تتألف منها بطارية ونج وهى : تحليل التآلفات ٠ تمييز الأصوات، ذاكرة الأصوات ، التوافق الموسيقى (الهارمونى) ، الشدة ، الايقاع ، الجمل الموسيقية ٠ وتتطلب الاختبارات الثلاث الأولى تمييزا حسيا على مستوى اكثر تعقدا من اختبارات سيشور ، أما فى الاختبارات الأربع الأخرى فانّ المفحوص يقارن النوعية الجمالية ٠ مقطوعتين موسيقيتين ، ومعنى ذلك ان بطارية ونج تؤكد جانب التذوق الموسيقى ٠ ويقترح ونج استخدام الاختبارات الثلاث الأولى كبطارية مختصرة مع الأطفال الصغار أو كأداة تصفية مبدئية ٠

وقد نشرت آمال أحمد مختار صادق (٢) فى عام ١٩٧٩ تقريرها عن تقنين هذا الاختبار على البيئة المصرية ٠ وقد بلغت عينة التقنين ٣٤١ مفحوصا من الذكور والاناث اشتقوا من مستويات تعليمية موسيقية مختلفة من دور المعلمين والمعلمات والمعهد القومى للموسيقى (الكونسيرفتوار).

وكلية التربية الموسيقية بجامعة حلوان - وقد سارت الباحثة فى تقدير صدق المحتوى من خلال استطلاع رأى المهتمين بالموسيقى بدرجة « الموسيقية » فى هذه الاختبارات · أما بالنسبة لصدق التكوين الفرض بحساب معاملات الارتباط بين الاختبارات الفرعية للبطارية وتحليل الفروق من مستويات التدريب الموسيقى وتمايز المستويات التعليمية الموسيقية ، كما استخدم الصدق المرتبط بالمحكات باستخدام بطارية سيشور التى سبق تقنينها من ناحية والتحصيل الموسيقى من ناحية أخرى ، وكانت هذه النتائج جميعا مدعمة لصدق البطارية ·

أما الثبات فقد حسب بطريقتين : التجزئة النصفية ، كيودر - ريتشارد سون · وحسبت المعايير المئينية للاختبارات الفرعية ولمجموع الاختبارات الادراكية ومجموع الاختبارات التذوقية والدرجة الكلية فى البطارية · كما حسبت أيضا نسب الذكاء الانحرافية التى اطلقت عليها الباحثة تسمية « نسب الذكاء الموسيقى » ومصطلح « الذكاء الموسيقى » هو المصطلح الأثير عند ونج مؤلف هذه الاختبارات ·

وفى عام ١٩٧٩ نشرت آمال صادق أيضا دراسة هامة تتضمن عرضا مفصلا لبطارية جديدة لقياس القدرات الموسيقية قامت بانشائها (٣) فى محاولة منها للربط بين مجال قياس القدرة الموسيقية والتقدم فى مجال دراسة القدرات العقلية بصفة عامة وما أحدثته نتائج البحوث فيها من تغير لم يظهر له أثر فى بطاريات سيشور وونج وبنتلى التى سبق الاشارة اليها ·

وتشمل الاختبارات الجديدة التى قامت الباحثة ببنائها ما يلى :

(١) اختبار الاغلاق الموسيقى : musical closure ويتألف من قسمين أولهما من ٢٤ مفردة فى كل منها يستمع المفحوص الى جملة موسيقية قصيرة ، وبعدها بفترة يستمع الى جملة موسيقية أكبر ، ويطلب منه تحديد ما اذا كانت الجملة الأولى متضمنة فى الجملة الثانية أم لا · وثانيهما ويتألف من ١٢ مفردة ، كل منها عبارة عن جملة موسيقية اختيرت من مؤلفات للبيانو ، ويعرض على المفحوص (موتيف) وعليه أن يقرر ما اذا كان موجودا فى اللحن أم لا ·

(٢) اختبار تعيين الألحان : melody identification ويتألف من صورتين احداهما للمراهقين والراشدين والثانية للاطفال · وتتكون صورة الكبار من قسمين أولهما يتألف من ٤٨ مفردة من الألحان في السلالم الكبيرة والصغيرة الغربية ، وثانيهما من ١٢ مفردة من الألحان التي تتضمن المسافة العربية (أي تمثل الموسيقى العربية) · أما اختبار الأطفال فيتألف من ٣٠ مفردة كلها من نوع الموسيقى الغربية · وفي جميع الحالات يقوم المفحوص بالمقارنة بين المعزوف والمدون على هيئة نوتة موسيقية ·

(٣) اختبار التعيين الايقاعي : rhythmic identification وقد أعد من صورتين أيضا · وتتألف صورة المراهقين والراشدين من ٦٠ مفردة ، عشرة منها في كل ميزان من الموازين الموسيقية ويطلب من المفحوص أن يقارن نمطا ايقاعيا منقرا بنمط ايقاعي مكتوب على هيئة تدوين موسيقي ، أما صورة الأطفال فتتألف من ٣٠ مفردة كلها في الميزان الموسيقي البسيط ·

(٤) اختبار التعرف على المسافات : interval recognition ويتألف من صورة واحدة تصلح للاطفال لسهولته الظاهرة بالنسبة للمراهقين والراشدين ، وينقسم الى قسمين : أولهما يتكون من ٣٠ زوجا من المسافات تعزف عزفا تتابعيا على البيانو باستخدام طبقات صوتية مختلفة ، ويطلب من المفحوص أن يقرر ما اذا كانت المسافتان متشابهتين أو مختلفتين · وفي القسم الثاني ٠٠ مفردات أكثر صعوبة تعزف عزفا تتابعيا على الكمان ·

(٥) اختبار تقدير المسافات : interval estimation ويتألف من صورة واحدة للمراهقين والراشدين لصعوبته الظاهرة بالنسبة للاطفال ويتألف من ٤٠ مفردة مصنفة الى ثلاثة اقسام تتناول على التوالي المسافات الصاعدة في الموسيقى الغربية ، والمسافات الهابطة منها ، ومسافات الموسيقى العربية · وفي جميع الأقسام تعزف في المفردة الواحدة مسافة مرجعية أصلية مرة واحدة على البيانو أو الكمان وبعد فترة توقف طولها ٥ ثوان تعزف ثلاث مسافات على التوالي ويطلب من المفحوص أن يحكم على أي هذه المسافات الثلاث هي الأقرب الى المسافة الأصلية ·

وقد قننت البطارية على عينة مصرية تتألف من تلاميذ المرحلة الاعدادية العامة (ن = ١١١) ، وتلاميذ المرحلتين الابتدائية والاعدادية بالمعهد القومي

للموسيقى (ن = ٧٤) ، وطلاب الكليات والمعاهد الموسيقية (ن = ٦٠) ٠ وحسب الثبات بطريقة كيودر ـ ريتشارد سون ، واستخدمت فى حساب الصدق محكات تتمثل فى الاختبارات التى تقيس القدرة الموسيقية والتى سبق تقنينها على البيئة المصرية (سيشور ، بنتلى ، ونج) كما استخدم محك التحصيل الموسيقى ، وكانت المؤشرات الاحصائية لكل من الثبات والصدق فى صالح البطارية ٠ كما استخدمت الباحثة طريقة أخرى فى حساب الصدق تتمثل فى المقارنة بين العينات ذات المستويات المختلفة من الانتقائية من كل من الأطفال والمراهقين والراشدين وكانت النتائج فى معظم الحالات دالة لصالح المجموعات الأعلى انتقائية (تلاميذ الكونسيرفتوار من الاطفال وطلاب المعاهد والكليات الموسيقية من المراهقين والراشدين) ٠

اختبارات القدرة فى الفنون البصرية :

توجد فى الوقت الحاضر طائفة ن الاختبارات التى تقيس بعض جوانب القدرة فى ميدان الفنون البصرية ، ومنها اختبار Graves Design Judgment Test الذى يهدف الى قياس قدرة المفحوص على ادراك المبادىء الأساسية للعمل الجمالى والاستجابة لها وهى مبادىء الوحدة unity والسيادة dominance والتنوع variety والتوازن balance والاستمرار continuity والتناسق symmetry والتناسب proportion والايقاع rhythm ٠ وقد تحقق ذلك بعرض ٩٠ ثنائية وثلاثية من الرسوم المجردة ، فى كل منها واحد من الرسوم يتفق مع المبادىء الثمانى هذه بينما يتمثل أحد هذه المبادىء او بعضها فى الرسم الآخر او الرسوم الأخرى ٠ ويطلب من المفحوص أن يفاضل بين رسوم الثنائية أو الثلاثية ويختار أحدها على اساس التفضيل الفنى ٠ ولم يتضمن هذا الاختبار نماذج من أعمال الفن ، لأن الحكم على الصور واللوحات يتأثر بالخبرة السابقة وسمات الشخصية والانفعالات ومفهوم المفحوص عن الفن٠ ولذلك فان الاختبار يؤكد محض الاختبار الجمالى ، أو ما يسميه أحد مؤلفى هذا الكاب « الحساسية الجمالية » (٢٠) ٠ وبالاضافة الى ذلك فان الاختبار يقيس العمليات الادراكية المعقدة وليس العناصر الحسية البسيطة التى تتدخل فى الحكم الجمالى ، وفى هذا يتشابه مع اختبارات ونج فى الموسيقى ، ومع المبادىء المتضمنة فى قياس العامل العام فى الذكاء ٠

والاختبار الثانى فى هذا المجال هو اختبار ماير للحكم الفنى ·
Meier Art Judgment Test

ويهدف الى قياس الحكم الجمالى بطريقة « كليـــة » · ويختلف عن اختبار جريفز فى انه يتألف من ١٠٠ زوج من اللوحات الفنية مرســـومة بالأبيض والأسود ، ويمثل أحد عنصرى الثنائية عملا فنيا ممتازا ، والعنصر الآخر أدخل عليه بعض التغيير والتعديل فى أحد الجوانب الفنية الهامة تجعله أقل جودة من العنصر الأول · ويحدد المفحوص أى الجوانب قد تغير (الشكل أو الزوايا مثلا) الا أنه لا يعلم أى اللوحتين هى اللوحة الأصلية ، ويطلب من المفحوص أن يحدد تفضيله لأى منهما ·

ويرى ماير أن هذا الاختبار يقيس الحكم الجمالى ، الا أنه فى رأى أحد مؤلفى هذا الكتاب يقيس « الحساسية الجمالية » أيضا (٢٠ : ٥) ، والفرق بينه وبين اختبار جريفز فرق فى المحتوى وليس فى العمليات الأساسية المتضمنة ·

واذا كان اختبار جريفز وماير ينتميان الى ميـدان التذوق الفنى فان اختبارى كنوبر Knauber وهـــورن Horn ينتميان الى ميدان الابتكار الفنى · ويسمى أولهما Knauber Art Ability Test ويصلح للمراهقين والراشدين ، ويتطلب من المفحوص ما يأتى :

١ ـ أن يرسم من الذاكرة أشكالا فى حدود المكان المسموح به ·

٢ ـ أن يرسم شخصيات نمطية (بابا نويل مثلا) ·

٣ ـ ترتيب مؤلف معين فى حدود المكان المسموح به ·

٤ ـ ابتكار وتكملة رسوم معينة من عناصر معطاة ·

٥ ـ تحديد الأخطاء فى أعمال مرسومة (مثل المنظور الخاطئ) ·

٦ ـ انتاج أعمال فنية تدل على الخيال والحذق والتمثيل الرمزى ·

٩ أما اختبار هورن ويسمى Horn Art Aptitude Inventory فقد أعد خصيصا للاستخدام فى أغراض قبول طلاب كليات الفنون الجميلة ، ويقيس على وجه الخصوص الابتكارية فى الفن مستخدما فى كل سؤال وصفا بسيطا مرنا أو مرشدا لكل رسم . ويتالف الاختبار من قسمين : الأول يتطلب من المفحوص أن يرسم ٢٠ شيئا مألوفا (منزل - كتاب ، الخ) فى وقت قصير ، كما يتطلب منه انتاج رسوم تجريدية بسيطة مبتدئا بمجموعة من المثلثات والمربعات وغيرها . أما القسم الثانى فيتطلب مستوى أعلى من الابتكارية وفيه يعطى للمفحوص ٢٠ بطاقة مستطيلة على كل منها خطوط تفيد كبداية لصورة ينتجها المفحوص . ويحكم على الرسوم على أساس الخيال الابتكارى من ناحية والخصائص الفنية من ناحية أخرى .

ومن الاختبارات المصرية التى تقيس القدرات فى الفنون البصرية اختبارات تكميل الأشكال وتكميل الصور والتقدير الجمالى التى أعدها الدكتور محمد عماد الدين اسماعيل (٢٥) لقياس عاملى الطلاقة والتقدير الجمالى . ويقيس اختبار تكميل الأشكال عامل الطلاقة وفيه يطلب من المفحوص أن يضيف الى خطوط مرسومة بضعة خطوط اخرى بحيث تجعل منه صورة لشىء مألوف . ويقيس اختبار تكميل الصور عامل الطلاقة ايضا وفيه يطلب من المفحوص أن يكتب فى مكان معين من صورة مرسومة اكبر عدد من الاشياء التى يمكن رسمها . أما اختبار التقدير الجمالى فهو تقنين مصرى لاختبار ماير الذى أشرنا اليه . وقد حسب صدق الاختبارات بايجاد معاملات الارتباط بين كل منها وبين المحك (وهو تقدير المدرسين للقدرة الفنية) فكانت النتائج كالآتى : ٤٤ر لاختبار تكميل الأشكال ، ٣٤ر لاختبار تكميل الصور . ٤٤ر لاختبار التقدير الجمالى (ماير) . أما عن الثبات فقد حسب بمعادلة كيودر - ريتشارد سون وكان معامل ثبات اختبار تكميل الأشكال ٥٥ر ، واختبار تكميل الصور ٨٣ر ، واختبار (ماير) للتقدير الجمالى ٥٤ر.

وحسبت معايير الدرجات التائية للاختبارات الثلاثة ، وكانت عينة التقنين فى كل الأحوال ٦٥ طالبا من المعهد العالى للتربية الفنية .

اختبارات التفكير الابتكارى :

من أهم خصائص التقويم النفسى المعـــــاصر زيادة الاهتمام بقياس الابتكارية مسايرة للنشاط المتزايد فى مختلف ميادين علم النفس لدراسـة طبيعة النشاط الابتكارى والعوامل المؤثرة فيه · ولا يتسع المقام للافاضة فى هذه المسائل، ويمكن الرجوع الى بعض المصادر المتاحة فى اللغة العربية (١٩) ونستطيع القول بايجاز أن الاهتمـــــام بالتفكير الابتكارى يعبر عن حاجة المجتمعات المعاصرة لزيادة وتنمية ثروتها البشرية من العلماء والمهندسـين والقادة الاداريين والفنانين · ويرى كثير من العلماء المهتمين بهذا الميدان أن الابتكار عملية أساسية فى هذه الأحوال جميعا ، فهو كما يقول تايلر « يختلف فى العمق والمجال وليس فى النوع » ومن رأيه عدم التمييز بين الابتكار العلمى والابتكار الفنى لأن الابتكار « يتضمن نظرة للمشكلات أكثر اساسية من التدريب المهنى » ·

ويعود الفضل الى العالم الأمريكي المعاصر جيلفورد فى زيادة حركة قياس التفكير الابتكارى · فقد أثرت نظريته عن تنظيم العقل فى معظم بحوث الابتكار ، وفيها يقترح تصنيف القدرات العقليـــــة الى فئتين كبيرتين هى قدرات التفكير ، وقدرات الذاكرة · وتنقسم عوامل التفكير الى ثلاثة انواع هى : قدرات التفكير المعرفى (بمعنى الاكتشاف) ، وقدرات التفكير الانتاجى وقدرات التفكير التقويمى · ثم تنقسم قدرات التفكير الانتــاجى الى فئتين فرعيتين هما : قدرات التفكير الانتاجى التقاربى ، وقدرات التفكير الانتاجى التباعدى · والانتاج التقاربى (أو الاختزالى) يتضمن تضييق الاحتمالات عند انتاج اجابة واحدة محتملة للمشكلة ، مثل : طويل الى قصير مثل مرتفع الى ··· ؟ وبالطبع فان أغلب اختبارات الذكاء العام تقيس هذا النمط من التفكير ، بل ان معظم الاختبارات العقلية التى أشرنا اليها فى الفصول السابقة من هذا الكتاب تنتمى الى فئة الانتاج التقاربى · أما التفكير التباعدى فيتطلب انتاج اكبر عدد ممكن من الاجابات · مثل : ما هى الاستخدامات التى يمكن أن تفكر فيها لقالب الطوب ؟ » ·

ويرى جيلفورد أن الانتاج التباعدى عامل هام فى التفكير الابتكارى · واستخدم فى وضع اختبارات عديدة لقياس هذه القدرة فى مستويات مختلفة

وفى محتويات متباينة · ومن رأيه أن بعض عوامل الذاكرة والتفكير المعرفى والتفكير التقويمى والتفكير الانتاجى التقاربى ترتبط بالعمـــل الابتكارى · ويعتقد على وجه الخصوص أن الموهبة الابتكارية يمكن قياسها بالاختبارات التى تقيس قدرات : الأصالة والطلاقة والمرونة من مختلف الأنواع ، وقدرات إعادة التحديد والتضمين والتفكير التقويمى أيضا ·

وقد أثرت نظرية جيلفورد فى ظهور بطاريتين رئيسيتين لقياس التفكير الابتكارى احداهما أعدها جيلفورد وزملاؤه بجامعة كاليفورنيا · والأخرى أعدها تورنس بجامعة مينيسوتا ·

اختبارات جيلفورد للتفكير الابتكارى :

تتألف بطارية اختبارات التفكير الابتكارى لجيلفورد على النحو الآتى :

١ ـ اختبارات جيلفورد ـ كريستنسن للطلاقة وتتألف من ٤ اختبارات

(أ) طلاقة الكلمات (الانتاج التباعدى لوحدات الرموز) : يطلب فيه من المفحوص أن يكتب كلمات ذات خصائص معينة ، كأن تتضـــمن حروفا معينة (ب مثـــلا) ·

(ب) طلاقة الأفكار (الانتاج التباعدى لوحدات المعانى) : يطلب فيه من المفحوص أن يذكر أسماء الأشـــياء التى تنتمى الى فئة معينة (مثل السوائل التى تحترق) ·

(جـ) طلاقة التداعى (الانتاج التباعدى للعلاقات بين المعانى) · يطلب فيه من المفحوص اعطاء كلمات تتشابه فى المعنى مع كلمة اخرى ·

(د) الطلاقة التعبيرية (الانتاج التباعدى لمنظومات الرموز) : يطلب فيه من المفحوص تكملة جمل ، كل منهـا يتألف من ٤ كلمات ، بحيث تبدأ كل كلمة بحرف معين ·

٢ ـ اختبار الاستخدامات البديلة (الانتاج التباعدى لفئات المعانى) :

يطلب فيه من المفحوص أن يذكر الاستخدامات المحتملة لشيء معين غير استخدامه المألوف (قالب الطوب مثلا يستخدم في البناء استخداما مألوفا) ·

٣ ـ اختبار المترتبات (الانتاج التباعدى لوحدات المعانى ، والانتاج التباعدى لتحويلات المعانى) · وفيه يطلب من المفحوص أن يذكر النتائج المختلفة المترتبة عن حادث افتراضى (مثل : ماذا يحدث لو أن الناس توقفوا عن النوم ؟) · ويعطى للمفحوص فى هذا الاختبار درجتان : احداهما للاستجابات الواضحة (وتقيس طلاقة الأفكار أو الانتاج التباعدى لوحدات المعانى) ، والأخرى للاستجابات البعيدة (وتقيس الأصالة ، والانتاج التباعدى لتحويلات المعانى) ·

٤ ـ الأعمال المحتملة (الانتاج التباعدى لتضمينات المعانى) · ويطلب فيه من المفحوص أن يذكر الأعمال المحتملة التى يمكن أن يرمز لها بشعار ما (مصباح كهربائى مثلا) ·

٥ ـ عمل الأشياء (الانتاج التباعدى لمنظومات الاشكال) : ويطلب فيه من المفحوص أن يرسم أشياء محددة باستخدام مجموعة معطاة من الأشكال (كالدوائر أو المثلثات) · وعلى المفحوص الا يستخدم الا الاشكال المعطاة فقط ·

٦ ـ الاسكتشات (الانتاج التباعدى لوحدات الأشكال) : ويشبه اختبار عمل الاشياء فيما عدا أن كل صفحة فى الاختبار تشتمل على نفس الشكل (دوائر مثلا فقط) ويطلب من المفحوص أن يرسم أكبر عدد من الاسكتشات باضافة تفاصيل الى كل شكل من الاشكال المعطاة ·

٧ ـ مشكلات عيدان الثقاب (الانتاج التباعدى لتحويلات الأشكال) : وفيه يطلب من المفحوص أن يستبعد عددا معينا من عيدان الثقاب (المرسومة على الورق فى صورة خطوط مستقيمة) بحيث ينتج عن ذلك عددا من المربعات أو المثلثات ·

٨ ـ اختبار الزخرفة (الانتاج التباعدى لتضمينات الاشكال) : وفيه

يطلب من المفحوص أن يزخرف صورا لأشياء مألوفة بأكبر عدد من الرسوم المختلفة .

ويلاحظ أن الاختبارات السبع الأولى من هذه القائمة تتطلب استجابات لفظية ، بينما تستخدم الاختبارات الأربع الأخيرة محتوى من الأشكال أو الصور . وهذه الاختبارات الأحد عشر تمثل عينة صغيرة من مجموعات الاختبارات التى أنتجتها بحوث جيلفورد وزملائه وتلاميذه بجامعة جنوب كاليفورنيا لأكثر من عشرين عاما . وقد أعدت معايير مبدئية لهذه الاختبارات فى صورة مئينيات ومعايير جيمية C-Score (وهو معيار يلخص المستويات التائية الى ١١ مستوى وهو فى ذلك يشبه المعيار التساعى ويختلف عنه فى عدد المستويات) . وفى معظم الأحوال اشتقت هذه المعايير لعينات من الراشدين أو طلاب من سن ١٥ + أو كليهما . وتصلح هذه الاختبارات بوجه عام لمستوى المرحلة الثانوية وما بعدها ، رغم أن جيلفورد يذكر أن بعض اختباراته تصلح لمستوى المرحلة الابتدائية (ابتداء من الصف الأول) .

والمشكلة الجوهرية فى اختبارات التفكير الابتكارى هى مشكلة التصحيح فهى تتطلب من المصحح تدريبا كافيا على اتباع تعليمات التصحيح ودراسة الأمثلة التى تتضمنها كراسات التعليمات . ورغم ذلك فان عملية تصحيح هذه الاختبارات تستغرق وقتا كبيرا .

وحسب جيلفورد ثبات هذه الاختبارات بطريقة التجزئة النصفية . وتتراوح معاملات الثبات بين ٦٠ر ، ٨٠ر . اما الصدق فقد اعتمد اعتمادا كبيرا على صدق التكوين الفرضى كما يتمثل فى اتفاق الاختبارات مع نموذج العقل الذى يقترحه . ولا زالت هذه الاختبارات فى حاجة الى انواع الصدق الأخرى وخاصة الصدق المرتبط بالمحكات .

وقد قام عبد السلام عبد الغفار (١١) باعداد بعض اختبارات جيلفورد باللغة العربية ، فأعد اختبارين لقياس الطلاقة اللفظية (طلاقة الكلمات) هما اختبار الطلاقة اللفظية ١ ، ويتطلب من المفحوص انتاج كلمات تبدأ

بحرف معين ، واختبار الطلاقة اللفظية ٢ ويتطلب من المفحوص أن ينتج كلمات تنتهي بحرف معين • كمـــا أعـــد ٣ اختبـــارات أخـــرى لقيـــاس بعض عوامل التفكير الابتكارى الأخرى وهي : اختبار الطلاقة الفكرية (لقياس عامل طلاقة الأفكار) ، واختيار الاستعمالات (لقياس عامل المرونة التلقائية، واختبار المترتبات (لقياس عامل الأصالة) • وقد حسب ثبات هذه الاختبارات من أداء عينة عددها ١٢٠ تلميذا بالمدارس الثانوية العامة بطريقة اعادة الاختبار لاختبارى الطلاقة اللفظية فبلغ معـــاملات الثبات ٦٢ر ، ٦٦ر • واستخدام في حساب ثبات اختبارات الطلاقة الفكرية والاستعمالات والمترتبات طريقة التجزئة النصفية وبلغت معاملات الثبات ٧٥ر ، ٦٩ر ، ٨٠ر ، على التوالى • وفى تقدير الصدق حسب الباحث معـاملات ارتبـــاط الاختبارات بالتحصيل المدرسى ، وتراوحت هذه المعاملات بين ٢٧ر (الاستعمالات) ٥٨ر (المترتبات) ، ولم تتضمن كراسة التعليمات بيانات عن معايير الاختبارات•

كما قام سيد خير الله (١٠) باعداد بعض اختبار جيلفورد أيضـــا وصدر له اختبار القدرة على التفكير الابتكارى يتكون من خمسة أقســـام تقيس : الاستعمالات والمترتبات ولعبة الكلمات ، ولا تتوافر لدينا أى معلومات سيكومترية عن هذا الاختبار •

اختبارات تورنس للتفكير الابتكارى :

أشرنا الى أن اختبارات جلفورد هى محصلة بحوثه العاملية فى ميدان التفكير الانسانى وطبيعة الذكاء ، أما اختبارات تورنس للتفكير الابتكـــارى التى نتناولها الآن فقد ظهرت لمواجهة الضرورات التربوية والعملية كجزء من برنامج طويل من البحوث تهتم بالخبرات التعليمية التى تساعد على تنميـــة الابتكار • وبعض اختبارات تورنس هى سى الواقـــع تعـــديل للاســـاليب التى استخدمها جيلفورد • بالاضافة الى أن الدرجات المشتقة من البطاريات الكاملة تعتمد على العوامل التى توصلت اليها بحوث جيلفورد وهى الطـلاقة والمرونة والأصالة واعطاء التفاصيل • وفى هذا الصدد يذكر تورنس أنه لم يسع الى وضع اختبارات « نقية عامليا » أو « بسيطة التكوين » ، وانما الى التوصل الى اختبارات تتضمن مواقف مماثلة لما تتطلبه العملية الابتكارية

فى الظروف الطبيعية والمعقدة • ولذلك فان كل اختبار قد يصحح فى ضوء عاملين ، أو ثلاثة عوامل ، أو العوامل الأربع جميعا •

وتتألف اختبارات تورنس للتفكير الابتكارى من ١٠ اختبارات ، مصنفة الى بطاريتين : احداهما لفظية والأخرى مصورة • وتسمى البطارية الأولى: التفكير الابتكارى بالكلمات ، والثانية : التفكير الابتكارى بالصور • وحتى يمكن التخفيف من التهديد الذى تتضمنه كلمة « اختبار » استخدمت فى كل الأحوال كلمة « نشاط » كما تؤكد التعليمات على « اللعب »، وتوصف هذه الاختبارات بانها تصلح لمختلف المستويات ابتداء من مرحلة الحضانة حتى مستوى الدراسات العليا ، بشرط ان تطبق فرديا وشفويا على المستويات الأدنى من الصف الرابع الابتدائى • وتوجد صورتان متكافئتان لكل بطارية من البطاريتين •

وفى « التفكير الابتكارى بالكلمات » الصورة ا والصورة ب تتألف الأنشطة الثلاث الأولى (اسأل وخمن) من صورة غير عادية يستجيب لها المفحوص بكتابه (أ) جميع الأسئلة التى يحتاج الى توجيهها لكى يعرف ما يحدث فى الصورة (ب) الأسباب المحتملة لما حدث فى الصورة (جـ) النتائج المترتبة على ما حدث فى الصورة • اما النشاط الرابع فيتعلق بطرق تحسين لعبة من لعب الأطفال بحيث تصبح مصدرا لمزيد من السرور والفرح لمن يلعب بها • ويتعلق النشاط الخامس بذكر الاستخدامات غير العادية لشىء مألوف : (علب الكرتون) فى الصورة (ا) ، (وعلب الصفيح) فى الصورة (ب) • ويتعلق النشاط السادس بالأسئلة غير المعتادة التى يمكن توجيهها عن الشىء المتضمن فى النشاط الخامس • اما النشاط السابع فقد اعد على نسق اختبار المترتبات عند جيلفورد ، ويتطلب من المفحوص ان يذكر كل ما يمكن أن يحدث اذا نشأ موقف غير ممكن الحدوث • وتعطى البطارية درجات كلية فى العوامل الثلاث : الطلاقة والمرونة والأصالة •

ويتألف اختبار « التفكير الابتكاري بالصور » : الصورة (أ) والصورة (ب) ، من ثلاثة انشطة :

(١) تكوين الصورة وفيه يعطى للمفحوص ورقة ملونة ذات شكل منحنى وعليه ان يضعها على صفحة بيضاء فى الموضع الذى يشاء ويستخدمها

كبداية لرسم صورة غير عادية (والورقة الملونة المستخدمة فى الصورة (أ) زرقاء اللون ، والمستخدمة فى الصورة (ب) صفراء اللون فى الطبعة العـــربية ٠

ويحصل المفحوص فى البطارية المصورة على ٤ درجات هى : الطلاقة والمرونة والأصالة والتفاصيل ٠

وقد أعد هذه البطاريات باللغة العربية (بموافقة مؤلفها تورنس) عبد الله محمود سليمان وفؤاد أبو حطب (٦) وظهرت عام ١٩٧١ على النحو الآتى :

١ ـ التفكير الابتكارى بالكلمات : الصورة ا ٠
٢ ـ التفكير الابتكارى بالكلمات : الصورة ب ٠
٣ ـ التفكير الابتكارى بالصور : الصورة ا ٠
٤ ـ التفكير الابتكارى بالصور ٠ الصورة ب ٠
٥ ـ كراسة تعليمات اختبارات تورنس للتفكير الابتكارى ٠
٦ ـ استمارة تصحيح (١) لاختبارات التفكير الابتكـــارى بالكلمات (الصورة (ا) أو الصورة (ب)) ٠
٧ ـ استمارة تصحيح (٢) لإختبارات التفكير الابتكارى بالصـــور (الصورة (ا) والصورة (ب)) ٠
٨ ـ استمارة تقدير المعلمين للابتكار ٠
٩ ـ استمارة تقدير التلاميذ للابتكار ٠

وفى عام ١٩٧٧ نشر معدا الاختبار تقريرهما الأول (١٦) عن تقنين هذه البطارية ، ويشمل على وجه الخصوص تقنين الصورة ا من اختبارات الأشكال أو الصور (*)، فطبق الاختبار على٣١٣ تلميذا وتلميذة منمدرستين اعداديتين بالقاهرة تتراوح أعمارهم بين ١٢ ، ١٦ سنة ، ولحساب الصدق استخدمت

(*) عاون الباحثين فى تصحيح الاختبارات التى طبقت فى البحث المشار اليـــه يومئذ السادة (الدكاترة فيما بعد) محمود السيد عمر ، على السيد سليمان ، سـيد محمود الطواب ، محمد متولى غنيمة ، والاستاذ محمد السعيد عبد الحليم ٠ وبعد انقضاء اكثر من ١٥ عاما على مشاركتهم الكريمة نزجى اليهـم الشـــكر والاعتراف بالفضـــل ٠

استمارة تقدير المعلمين للابتكار حيث طلب من مدرسى الفرق التى أجريت عليها الاختبارات أن يختار كل منهم ٥ تلاميذ يقعون فى الفئة العليا ، ٥ تلاميذ يقعون فى الفئة الدنيا فى كل من أبعاد الابتكار المقيسة وهى : الطلاقة ، المرونة ، الأصالة ، التفاصيل · واختير كل تلميذ يتفق عليه مدرسان على الأقل · وحسبت دلالة الفروق بين متوسطات مجموعات التلاميذ الأعلى والأدنى فى مختلف ابعاد الاختبار باستخدام مقياس (ت) فوجدت الفروق فى معظم الأحوال (باستثناء التفاصيل) دالة وفى الاتجاه الصحيح الا أن مستوى الدلالة كان عاليا (١٠ر) بالنسبة الى بعد المرونة ، ومنخفضا نسبيا (١٠ر) بالنسبة الى بعدى الطلاقة والأصالة ·

وقد استخدم فى حساب الثبات الطرق الآتية ·

١ _ حساب معاملات الارتباط بين متغيرات الاختبار (الطلاقة ، المرونة ، الأصالة ، التفاصيل ، الدرجة الكلية) فوجدت المعاملات فى كل الأحوال عالية وذات دلالة احصائية ·

٢ _ حساب معاملات الارتباط بين مصححين قاما بتصحيح ٣٨ اختبارا كل على حدة ، وكانت معاملات ثبات التصحيح ٩٩ر ٠ للطلاقة ، ٩٩ر٠ للمرونة ، ٩٩ر للاصالة ، ٩٧ر٠ للتفاصيل ، ٩٩ر٠ للدرجة الكلية · ويرجع ارتفاع المعاملات على هذا النحو الى التزام المصححين بقواعد مقننة فى التصحيح ، كما لجأ الباحثان فى تدريب المصححين على طريقة تورنس ، وتأكدت بالفعل جدواها ·

٣ _ حساب الثبات بطريقة اعادة الاختبار (ن = ٣٨) وبلغت المعاملات ٥٤ر للطلاقة ، ٣٦ر٠ للمرونة ، ٥٣ر٠ ، للاصالة ، ٤٠ر٠ للتفاصيل ، ٥٠ر٠ للدرجة الكلية فى الاختبار ·

وقد أعد الباحثان معايير الاختبار فى صورة درجات تائية ، كما أعدا النماذج التجريبية لتصحيح الصورة (أ) من اختبارات الأشكال ، وهى نماذج مشتقة من استجابات عينة التقنين المصرية ذاتها وليست منقولة عن كراسات التعليمات الأجنبية ·

ولسوء الحظ لم تتوافر لمعدى الاختبار باللغة العربية فرصة اخرى للعكوف على الاختبارات الثلاثة الأخرى فى البطارية لتقنينها ، كما لم تتح لهما الفرصة لاعادة تقنين اختبار الأشكال (الصورة ١) الذى مر على تقنينه الأول الآن أكثر من خمس عشرة سنة، الا ان مما يخفف من وقع الأمر ان الاختبارات الأربعة استخدمت فى بعض البحوث الامبريقية وفى عدد من رسائل الماجستير والدكتوراه فى الجامعات المصرية وعدد من جامعات العالم العربى تؤكد نتائجها صلاحيتها للاستخدام ٠

اختبار التفكير الابتكارى للاطفال :

اعدت هذا الاختبار فى أصله الأجنبى سيلفيا ريم S. Rimm عام ١٩٧٦ بهدف قياس التفكير الابتكارى عند الأطفال ويطلق عليه اسم بطارية الاختبارات الجماعية للكشف عن الموهوبين Group Inventory for Finding Talent (GIFT)

ويتكون من جزئين رئيسيين : أولهما ويصلح للاطفال من سن ٦ ـ ٩ سنوات ، وثانيهما ويصلح للاطفال من سن ٩ ـ ١٢ سنة ويتضمن جزءا مصورا من اختبارات تورنس (الدوائر والمربعات) وجزءا لفظيا من اختبارات جيلفورد (الاستعمالات) ، بالاضافة الى جزء يتألف من عبارات تتضمن صفات الطفل المبتكر ٠

وقد اعد الاختبار باللغة العربية سيد خير الله ومحمود عبد الحليم منسى (١١) ، وقاما بتقنينه على عينة من ٥٠٠ طفل من الذكور والاناث من تلاميذ المدارس الابتدائية بمدينة الاسكندرية ٠ وحسب الثبات بطريقة اعادة الاختبار وتراوحت معاملاته لمختلف صفوف المرحلة الابتدائية بين ٥١ر ، ٨٣ر ٠ وفى حساب الصدق استخدم الباحثان تقديرات المعلمين للتفكير الابتكارى عند تلاميذهم كمحك ، كما استخدم أيضا اختبار القدرة على التفكير الابتكارى من اعداد سيد خير الله كمحك آخر ٠

اختبار التفكير الابتكارى عند الأطفال باستخدام الحركات والأفعال :

اعد هذا الاختبار فى أصله الاجنبى تورنس Torrance وجبس

Gibbs عام ١٩٧٧ ويعتمد فى جوهره على استخدام الحركات والأفعال ليكون مناسبا للاطفال بين ٣ - ٧ سنوات· ويتناول اربعة أنشطة على وجه الخصوص هى : كم طريقة ؟ ، ونقدر نتحرك مثل ؟ ، هل فيه طريقة أخرى ؟ الاستعمالات · وتقدر لهذه الأنشطة درجات للطلاقة والتخيل والأصالة ·

وقد أعد الاختبار باللغة العربية محمد ثابت على الدين (٢٣) ، وأعد له معايير تصحيح ملائمة للعينات المصرية · وطبقه على مجموعات من الأطفال المقيدين بدور الحضانة وبعض المدارس الابتدائية بالمنصورة عام ١٩٨١ ، وحسب له الثبات بطريقة اعادة الاختبار · وكانت المعاملات مرضية، وبذل الباحث محاولات جادة للعثور على محكات يحدد فى ضوئها صدق الاختبار ، ولم يصلح استخدام الرسم كمحك ، كما لم تصلح تقديرات المشرفين المتصلين بأطفال الحضانة ، وكان أنسب المحكات هو تمايز العمر ·

اختبارات واطسون وجليزر للتفكير الناقد :

أعد هذا الاختبار جودون واطسون وادوارد جليزر ليزود المفحوص بمجموعة من المواقف والمشكلات التى تتطلب استخدام التفكير الناقد ، ويتكون من ٥ اختبارات فرعية :

١ - الاستنتاج : ويقيس القدرة على التمييز بين الدرجات المختلفة من الصواب والخطأ ، والتوصل الى استنتاجات معينة على أساس البيانات المعطاة ·

٢ - التعرف على الافتراضات : ويقيس القدرة على التعرف على الافتراضات المتضمنة فى القضايا المعطاة ·

٣ - الاستنباط : ويقيس القدرة على التفكير الاستنباطى من مقدمات معينة ·

٤ - التفسير : ويقيس القدرة على الحكم على الشواهد والأدلة ، وللتمييز بين التعميمات التى تبررها الأدلة والتى لا تبررها ·

٥ - تقويم الحجج : ويقيس القدرة على التمييز بين الحجج القــوية والضـعيفة بالنسبة لمسائل مطروحة أمام المفحوص ٠

وقد أعد هذا الاختبار فى صورته العربية جابر عبد الحميد جابر ويحيى حامد هندام (٩) ٠

وقد استخدم الاختبار فى عدد من البحـــوث ورســـائل الماجستير والدكتوراه التى تناولت التفكير الناقد وتوافرت عنه بيانات سيكومترية كافيـــة ٠

بطارية اختبارات الاستعدادات الحسية الحركية للمكفوفين :

أعد هذه البطارية فرج عبد القادر (٢٢) لقياس الاستعدادات الحسية الحركية ، للمكفوفين ، وتتألف من الاختبارات الآتية :

١ - اختبار تمييز اللمس
٢ - اختبار تمييز الأبعاد ٠
٣ - اختبار تمييز الوزن ٠
٤ - اختبار المهارة اليدوية ٠
٥ - اختبار مهارة الأصابع ٠
٦ - اختبار تآزر اليدين ٠
٧ - اختبار تصنيف الشكل - اللمس ٠

ولحساب الثبات طبقت الاختبارات على عينة من ٣٥ من المكفوفين وأعيد التطبيق عليهم بفاصل زمنى لا يقل عن شهر وكانت معاملات الارتباط دالة فيما عدا اختبار مهارة الاصبع ٠ وحسب صدق التكوين الفرضى فى ضــوء ما هو متوقع نظريا عن الفروق بين الجنسين ، واستخدمت معايير العمـــر العقلى والدرجات المئينية والدرجات المعيارية ٠

مقياس التذكر المباشر :

من اعداد وفية الكمونى (٣٠) فى رسالتها للدكتوراه عام ١٩٧٦ ويتكون من اختبار للتعرف واختبار الاسترجاع بواسطة مثير مساعد ، واختبار الاسترجاع الحر المباشر واختبار سعة الأرقام • وحددت الصدق كما تقول فى ضوء « التقدير الشخصى أو الاعتماد على آراء الحكام والخبراء المشهود بكفاءتهم » ، وهى طريقة لا يعتد بها كثيرا فى حساب الصدق • قد تفيد فى المراحل الأولى لبناء الاختبار ، انما لا تعد طريقة كافية تصل بالاختبار الى درجة من الصلاحية يعتمد عليها ، كما أن الطريقة التى شرحتها الباحثة فى حساب الثبات (٣٠ : ٤ _ ٧) غير واضحة • وعلى هذا لا نوصى باستخدام هذا الاختبار بسبب ضعف معالمه السيكومترية •

اختبار الأشكال المتضمنة :

هذا الاختبار من اعداد وتكن Witkin وزملائه لقياس أحد الأساليب المعرفية التى شاعت فى السنوات الأخيرة وهو بعد الاعتماد على المجال فى مقابل الاستقلال عنه •

وقد نقل هذا الاختبار الى العربية أنور الشرقاوى وسليمان الخضرى الشيخ (٨) وحسب ثباته بطريقة التجزئة النصفية من أداء عينة من طلاب جامعة الزقازيق (ن = ١٦٥) وبلغ معامل الثبات للبنين ٧٦ر ، وللبنات ٧٨ر• كما استخدمت طريقة جتمان فى حساب الثبات وتكاد تكون القيم الاحصائية متطابقة فى الحالتين •

واعدت للاختبار معايير فى صورة السباعى المعيارى ، وقد أجريت على الاختبارات بحوث كثيرة تعد نتائجها شواهد على صدق التكوين الفرضى له•

مراجع الباب الرابع

١ ـ احمد زكى صالح : اختبار القدرات العقلية الأولية · مكتبة النهضة المصرية (ب·ت) ·

١ ا ـ اسحق حنا بطرس : مدى التوافق بين اختيار كليات الهندسة والقدرات اللازمة للنجاح منها ، رسالة ماجستير ، كلية التربية جامعة عين شمس ، ١٩٧١ ·

٢ ـ آمال أحمد مختار صادق : تقنين اختبارات ونج للذكاء الموسيقى على البيئة المصرية فى : بحوث فى تقنين الاختبارات النفسية (تحرير فؤاد أبو حطب) المجلد الثانى · مكتبة الأنجلو المصرية ، ١٩٧٩ ·

٣ ـ ــــــــــ : مقاييس جديدة للقدرة الموسيقية ، فى المرجع نفسه ·

٤ ـ ــــــــــ : اعداد صورة مختصرة من بطارية سيشور ، فى المرجع نفسه ·

٥ ـ آمال أحمد مختار صادق : كراسة تعليمات اختبارات سيشور للقدرات الموسيقية · مكتبة الأنجلو المصرية ، ١٩٧٢ ·

٦ ـ آمال أحمد مختار صادق : كراسة تعليمات اختبارات القدرات الموسيقية للاطفال « بنتلى » مكتبة الأنجلو المصرية ، ١٩٧٢ ·

٧ ـ ــــــــــ ، اكرام مطر : القيمة التنبؤية لاختبارات القبول للتعليم الموسيقى فى مصر · المجلة الاجتماعية القومية ، المجلد ١٣ ، العدد ٣ ، سبتمبر ١٩٧٦ ·

٨ ـ أنور محمد الشرقاوى ، سليمان الخضرى الشيخ : اختبار الأشكال المتضمنة · مكتبة الانجلو المصرية ١٩٨٥ ·

٩ ـ جابر عبد الحميد جابر ، يحيى حامد هندام : كراسة تعليمات اختبارات التفكير الناقد · دار النهضة العربية ، ١٩٧٠ ·

١٠ ـ سيد محمد خير الله : اختبار القدرة على التفكير الابتكارى (ب٠ت)

١١ ـ سيد محمد خير الله ، محمود عبد الحليم منسى : اختبار التفكير الابتكارى للاطفال (ب٠ت) ٠

١٢ ـ السيد محمد خيرى ٠ علم النفس الصناعى : دار النهضة العربية ١٩٦٧ ٠

١٣ ـ السيد محمد خيرى : بطارية اختبارات استعدادات حرف المعادن مصلحة الكفاية الانتاجية والتدريب المهنى (ب ٠ ت) ٠

١٤ ـ عبد السلام عبد الغفار : اختبارات القدرة على التفكير الابتكارى ٠ مطبعة البيان العربى ، ١٩٦٥ ٠

١٥ ـ عبد الله سليمان ، فؤاد أبو حطب : كراسة تعليمات اختبارات تورنس للتفكير الابتكارى ٠ مكتبة الأنجلو المصرية ، ١٩٧١ ٠

١٦ ـ عبد الله محمود سليمان ، فؤاد أبو حطب : تقنين اختبارات تورنس للتفكير الابتكارى على البيئة المصرية ٠ اختبارات الأشكال (الصورة ا) ٠ فى بحوث فى تقنين الاختبارات النفسية ، (تحرير فؤاد أبو حطب) ، المجلد الأول ، مكتبة الأنجلو المصرية ، ١٩٧٧ ٠

١٧ ـ عبد الوهاب محمد كامل : اختبار هيبكى نبراسكا لقياس الاستعداد للتعلم ، طنطا : المكتبة القومية الحديثة (ب ٠ ت) ٠

١٨ ـ على حسين بدارى ، أنور رياض عبد الرحيم : الاختبارات المعرفية مرجعية العوامل (غير منشور)

١٩ ـ فؤاد أبو حطب ، القدرات العقلية ٠ مكتبة الأنجلو المصرية (ط ٤) ، ١٩٨٣ ٠

٢٠ ـ ــــــــــــــ ، التفضيل الفنى وسمات الشخصية ، المجلة الاجتماعية القومية ، يناير ١٩٧٣ .

٢١ ـ ــــــــــــــ ، آمال صادق: كراسة تعليمات اختبار مينيسوتا للوحات الأشكال (غير منشور) .

٢١ أ ـ فتحى الزيات : تقنين بطارية اختبارات فلاناجان على طلاب كلية الطب . رسالة ماجستير ، كلية التربية جامعة المنصورة ، ١٩٧٨ .

٢٢ ـ فرج عبد القادر طه : اختبارات الاستعدادات الحسية الحركية للمكفوفين . مطبعة دار التأليف ، ١٩٧٤ .

٢٣ ـ محمد ثابت على الدين : اختبارات تورنس للتفكير الابتكارى عند الأطفال باستخدام الحركات والافعال . مكتبة الأنجلو المصرية ، ١٩٨٥

٢٤ ـ محمد عبد السلام أحمد . القدرة الكتابية واختباران لقياسها : لجنة التأليف والترجمة والنشر ، ١٩٦٠ .

٢٥ ـ محمد عماد الدين اسماعيل . قياس الابداع الفنى وعلاقته بالتربية الفنية ، النهضة المصرية ، ١٩٥٩ .

٢٦ ـ محمد عماد الدين اسماعيل ، سيد عبد الحميد مرسى : كراسة تعليمات اختبارات المهن الكتابية . مكتبة النهضة المصرية (ب.ت)

٢٧ ـ محمود السيد أبو النيل : الحد الأدنى اللازم لأداء الاختبارات النفسية للسائقين . المركز القومى للبحوث الاجتماعية والجنائية ، ١٩٧٥ .

٢٨ ـ محمود عبد القادر : تقنين بطارية الاستعدادات العامة . المركز القومى للبحوث الاجتماعية والجنائية ، ١٩٧١ .

١٨ أ ـ مصطفى خليل الشرقاوى : دراسة لأنماط القدرات اللازمة للتكيف والنجاح فى التدريب على بعض المجموعات المهنية وتطبيقاتها فى مجال

التربية والتوجيه المهنى · رسالة دكتوراه ، كلية التربية جامعة عين شمس ، ١٩٧٢ · ادلة بطارية اختبارات الاستعدادات الفارقة ·

٢٩ ـ هدى برادة ، فاروق صادق : بطارية اختبـــارات القدرات النفسية اللغوية وتقنينها على عينة مصرية فى كتاب : بحـــوث فى تقنين الاختبارات النفسية (تحرير فؤاد ابو حطب) ، المجلد الثانى ، مكتبة الأنجلو المصرية ، ١٩٧٩ ·

٣٠ ـ وفية الكمونى : مقياس التذكر المباشر ومتغير المدى · مطبعة النهضة الجديدة ١٩٧٦ ·

31. Anastasi, A. Psychological Testing. Macmillan, 1982.
32. Bennett, G.K., et al. Test of Mechanical Comprehension, New York, Psychol. Corporation, 1940-1954.

33. Anastasi, A. Drake, J. An empirical comparison of certain techniques for estimating the reliability of speeded tests. Educ. psycho. Measment, 1954, 14, 529-540.

34. Cronbach, L.J., Essentials of Psychological testing, Harper, 1984.

35. Flanagan, J. C. Flanagan Aptitude classification Tests. Technical Report, Chicago, SRA, 1959.

36. Fleishman, E.A. Dimentional analysis of movement reactions, J. Exp. Psychol. 1958, 17, 522-532.

37. Freeman, F.S. Theory and Practice of Psychological testing. (3rd ed.), Holt, Rinehart, Winston, 1962.

38. Guilford, J.P. The nature of human intelligence, McGraw-Hill, 1967.

39. Likert, R., Quasha, W.H. Revised Minnesota Paper Formbord. Psychol. corporation, 1941-1984.

40. Mellenbruch, P.L. mechanical motivation test. Psycho. Affiliates, 1957.

41. Melton, H.W. (ed.) Apparatus Tests. Washington, D.C. Government printing offices, 1947.

42. Mourad S.A. & Moustafa M.M. The Development of an Arabic Aptitude test. A.U.C., Egypt.

43. Nunnally, J.C., (Jr.) **Introduction to psychological Measurement.** McGraw-Hill, 1970.

44. Richardson, B. et al. SRA Mechanical Aptitudes, Chicago; Science Research Associate, 1949-1950.

45. Super D.E. (ed.) **The use of multifactor tests in guidance.** Washington, Amer. Pers and Guidance Assoc., 1959.

46. Super, D.E., crites, J.O. **Appraising vocational fitness by means of psychological tests.** Harper, 1962.

البـــاب الخامس

اختبــــارات التحصيل

الفصل الحادى عشر

بناء الاختبار التحصـيلى

يقصد بالاختبار التحصيلى achievement test الأداة التى تستخدم فى قياس المعرفة والفهم والمهارة فى مادة دراســــية أو تدريبية معينة أو مجموعة من المواد ·

وهذه الاختبارات هى أكثر الأدوات شيوعا فى التقويم التربوى ، وهى فى بلادنا هى الوسيلة الوحيدة التى تستخدم فى توجيه التلاميذ وانتقائهم ، كما انها فى اغلب الأحوال ما يمكن للمعلم أن يصنعه فى ممارسته اليومية للتدريس · ولذلك فانها تحتل منزلة خاصة بين أدوات التقويم النفسى يجعل من المهم افراد هذا الفصل لها بحيث يمكن تناولها بشىء من التفصـيل حتى يمكن للمعلم ، والسلطات التربوية عامة ، أن تستفيد من المبادىء التى سوف نتناولها هنا فى اعداد هذه الاختبارات ·

تحديد الأهداف التعليمية :

أشرنا فى الفصل الأول من هذا الكتاب الى العلاقة بين التقويم النفسى والأهداف التربوية educational objectives · والواقع أن هذه الأهداف تتفاوت من العبارات الشاملة التى تدل على الأغراض العامة والقيم والفلسفات التربوية الى العبارات المحددة التى تبين أساليب أداء نوعية أو خاصة · ويشار الى النوع الثانى فى علم النفس التربوى المعاصر بالأهداف التعليمية instructional objectives ، واذا كانت الأهــــداف التربوية يعوزها الوضوح فى أغلب الأحوال ، فان الأهداف التعليمية أكثر تحديدا لأنها تتناول النواتج النهائية لعملية التعلم والتعليم فى صــــورة أداء يمكن ملاحظته ، وبالتالى فهى على درجة أكبر من الوضوح ·

ويرى كثير من علمــــاء النفس التربويين المعاصرين (١٠) أنه من

الواجب صياغة الأهداف التعليمية فى عبـــارات صريحة explicit تدل على الأداء النهائى terminal performance لثلاثة أسباب هى :

١ ـ يؤدى تحديد الأهداف التعليمية تحديدا سلوكيا بهذا الشكل الى توجيه المعلم فى تخطيط عملية التعليم فيمكنه ان يحدد الخطوات التى يجب أن يسير فيها التلميذ للوصول الى الهدف ، كما يمكن أن ينهض بأعباء جميع الاستجابات التى يجب أن تصدر عن التلميذ للوصـــول الى الاستجابات الهامة النهائية ٠

ومن المؤكد ان اساليب التدريس تعتمـــد على الأهـــداف التعليمية والمدخلات السلوكية ، ومع ذلك فان من نتائج تخطيط المعلم هذا وبحثه قد يكون تعديل الأهداف بتغيير مقدار الزمن مثلا أو المستوى المتوقع للاتقان او الموضوعات التى يجب تناولها فى التدريس ٠ وبالطبع لا يمكن لهـــذه التغيرات أن تحدث ما لم تتم صياغة الأهداف بصورة صريحة ٠ وفى نفس الوقت لا يستطيع المعلم ان يوائم بين اساليبه فى التعليم والفروق فى المدخلات السلوكية ما لم يحدد بشكل دقيق ما يستطيع التلميذ اداءه قبل التعليم وما يتوقع منه اداءه بعد التدريس ، ومعنى ذلك أن أهم أسباب وجوب صياغة الأهداف بشكل صريح ما يهيؤه ذلك للمعلم من ارشاد فى تخطيط أساليبه فى التدريس ٠

٢ ـ يفيد تحديد الأهداف التعليمية فى عبارات صريحة فى تقويم الأداء ومن المعروف أن الاهتمام بالصياغة الدقيقة للاهـــداف التعليمية بدأه فى الأصل العلماء المهتمون ببناء الاختبارات وتقويم المناهج المدرسية ، فقــد اكتشفوا أن استخدام العبارات الغامضة فى التعبير عن الأهداف يجعل عملية بناء الاختبار ومفرداته عملية صعبة أن لم تكن مستحيلة ٠

٣ ـ تفيد ترجمة الأهداف التعليمية فى صورة أداء نهائى فى توجيه انتباه التلميذ وجهوده ٠

وقد اثبت ميجر ومكان (١٩) أن التعلم يمكن أن يحدث حين لا يفعل

المعلم اكثر من اعطاء التلاميذ قائمة بالأهداف التعليمية الواضحة يقوم التلاميذ بالاهتداء بها فى تعلمهم الذاتى ٠

وصف العمل التربوى :

يتطلب تحديد الأهداف التعليمية وصفا دقيقا للعمل التربوى كمايتطلب تحليلا تفصيليا لهذا العمل ٠ وطريقتا وصف العمل task description وتحليل العمل task analysis نشأتا أول الأمر فى ميدان علم النفس الصناعى ثم انتقلنا فى السنوات الأخيرة الى ميدان علم النفس التربوى ٠

ومن اهم أساليب وصف العمل التربوى طريقة ميجر Mager وطريقة ميللر Miller ، وتتطلب طريقة ميجر (٧ ، ١١) فى وصف العمل ـ او بعبارة أخرى تحديد الأهداف التعليمية ـ خطوات ثلاثة هى :

١ ـ تحديد الأداء النهائى الذى يسعى التعليم او التدريس الى انتاجه ٠
٢ ـ وصف الظروف والأحوال الهامة التى نتوقع حدوث السلوك فيها ٠
٣ ـ وصف مستوى الجودة الذى يجب أن يصل اليه اداء التلميذ ٠

فمثلا قد يكون الهدف التعليمى الذى نسعى الى تحقيقه هو أن يستطيع التلميذ تحديد العوامل التى أدت الى حدوث الحرب العالمية الأولى ، وهذا هدف مترجم فى اداء يمكن ملاحظته ملاحظة صريحة وعامة فى صورة أداء نهائى يسعى التعليم الى انتاجه (الخطوة الأولى) ، ثم يمكن أن نعطى للتلميذ قائمة بالعوامل التى تؤدى الى الأحداث التاريخية الهامة ، وتكون هذه القائمة وصفا لما يسميه ميجر العوامل الهامة التى يحدث فى ظلها السلوك النهائى (الخطوة الثانية) ، وأخيرا تتحدد المستويات المطلوبة (الخطوة الثالثة) ٠ والواقع أن تحديد المستويات هذا ليس من الأمور القاطعة ويتراوح بين الاتقان الكامل وبين اقل مستويات الاتقان ٠

وتشير طريقة ميللر (١٠) فى وصف العمل الى وجود ثلاث خطوات ايضا هى :

١ ـ ظروف الاستثارة stimulus condition او ما يسميه indication .

٢ ـ التنشيط activation او الاستجابة التى يجب أن تصدر عن التلميذ ٠

٣ ـ التغذية الراجعة Feedback او تحديد مستوىملاءمة الاستجابة٠

والواقع أن هذه الخطوات لا تختلف كثيرا عن تلك التى اشار أليها ميجر ، فظروف الاستثارة هى ما يشير اليه ميجر بالعوامل او الظروف او الأحوال الهامة ، والتنشيط عند ميللر يقابل ما يسميه ميجر بالأداء النهائى ، ويتشابه مفهوم التغذية الراجعة عند الأول مع مفهوم مستوى الأداء عند الثانى ٠

تحليل العمل التربوى :

بعد ما يقوم المعلم بوصف الأعمال التى يتوقع من تلاميذه اداءها يكون عليه تحليل هذه الأعمال وتصنيفها ٠ والغرض الرئيسى من تحليل العمل هو مساعدة المعلم على تحديد أنسب ظروف التعلم ، والتى تختلف من نوع من انواع التعلم لآخر ٠ فتعلم تكوين المفاهيم او حل المشكلات او الابتكار يتطلب ظروفا تختلف عن اكتساب المهارات الحركية او حفظ قائمة من الكلمات ٠ وهذا يعنى بالطبع أن مسئولية المعلم اكبر من مجرد وصف العمل التربوى او صياغة الأهداف فى عبارات سلوكية صريحة ٠ فهى تشمل ايضا تصنيف هذه الأساليب السلوكية المتضمنة فى كل هدف من هذه الأهداف ثم تحديد انسب شروط التعلم لكل منها ٠ .

ويوجد فى الوقت الحاضر تصنيفان اساسيان للسلوك لهما اهميتهما التربوية هما تصنيف جانيه Gagné وتصنيف بلوم Bloom ٠ ويشمل تصنيف جانيه (١٤) ثمانية انواع من التعلم تتدرج من البسيط الى المركب هى:

١ ـ التعلم الاشارى Signal learning او ما يسمى فى سيكولوجية

التعلم بالاشتراط الكلاسيكى وفيه يتم تعلم استجابة شرطية لمثير او اشارة معينة ٠

٢ - تعلم المثير - الاستجابة أو ما يسمى بالاشتراط الاجرائى ٠

٣ - التسلسل chaining او ما يسمى اكتساب المهارات الحركية (غير اللفظية) ٠

٤ - الارتباط اللغوى Verbal association .
٥ - التمييز المتعدد Multiple discimination
٦ - تعلم المفاهيم Concept learning .
٧ - تعلم المبادىء Principle learning .
٨ - حل المشكلة Problem solving او التفكير ٠

اما التصنيف الذى وضعه بلوم Bloom ومساعدوه للميدان العقلى المعرفى (٩) فيشتمل على فئتين كبيرتين هما المعرفة Knowledge والقدرات والمهارات العقلية بحيث نحصل فى النهاية على ستة فئات للاهداف التربوية هي :

١ - المعرفة Knowledge .
٢ - الفهم Comprehension .
٣ - التطبيق Application .
٤ - التحليل Analysis .
٥ - التركيب Synthesis .
٦ - التقويم Evaluation .

وصف المدخلات السلوكية :

تتطلب عملية التعليم وصفا للمدخلات السلوكية لدى التلاميذ حتى يمكن للمعلم ان يخطط لتنفيذ عملية التدريس ، بل وحتى يمكن للسلطات التربوية ان تخطط للعملية التربوية عامة ويتطلب هذا الوصف تحديدا دقيقا

الموضع الراهن لمعارف التلميذ ومعلوماته ومهارته وطرق تفكيره وقدراته واستعداداته وميوله ودوافعه وخصائص شخصيته ، قبل أن نسعى الى اكسابه النماذج السلوكية الجديدة من خلال التعلم والتدريس · وهذه البيانات يجب استخدامها لأغراض الوصف وليس لأغراض التفسير ، لأن هذه المفاهيم تصف فى الواقع المستوى السلوكى للفرد فى لحظة معينة ، ولا تستطيع الاختبارات والمقاييس التى نستخدمها فى قياس هذه الظواهر أن تفسر لنا أداء الفرد فيها ، فلا يمكن مثلا أن نعزو ضعف الأداء فى اختبار للذكاء الى « ضعف الذكاء » لأن ذلك يصبح نوعا من تحصيل الحاصل ولا يزيد فهمنا للسلوك الانسانى ·

والواقع أن وصف المدخلات السلوكية ضرورى لأى برنامج تربوى · فعملية التعليم يجب أن تبدأ مما لدى التلاميذ ويجب أن تستمر الى حيث يستطيعون من أداء نهائى تحت شروط معينة ولمستوى معين · كما أن تحديد اهدافنا يمكننا من أن نقرر ما سوف نعلمه وما نفترضانه لدى التلاميذ بالفعل ويمكن للتفاعل بين الأهداف والمدخلات السلوكية أن يؤدى الى تغيير الأهداف اذا ثبت أنها غير واقعية ، كما تتفاعل المدخلات السلوكية مع وسائل التدريس وأساليبه ، فنحن لا نبدأ تدريسنا لتحقيق هدف تربوى معين بمواد تعليمية جديدة تماما وانما قد نستدعى الاستجابات التى سبق تعلمها وخاصة تلك التى تتناسب مع موقف التعلم الجديد ·

ثم ان لكل نوع من التعلم مدخلاته السلوكية فتنمية سلوك حل المشكلة تتطلب معرفة بمدخلات سلوكية تختلف عما يتطلبه تنمية احدى المهارات الحركية (٦) ·

العلاقة بين الأهداف التعليمية وأساليب التقويم التربوى :

يلخص جرونلاند (١٥) العلاقة بين الأهداف التربوية والأهداف التعليمية وأساليب التقويم التربوى على النحو الآتى :

الأهداف التربوية العامة « الأهداف التى توجه التدريس توجيها يوميا»



نواتج التعلم الخاصة أو النوعية « الأهداف التعليمية »

« سلوك التلاميذ واداؤهم الظاهر الذى يعد دليلا عن تحقيق هذه الأهداف»



اساليب التقويم

(الأساليب التى يحصل بها على عينة من سلوك التلاميذ كما ترصف فى نواتج التعليم النوعية) ٠

ويتطلب هذا ربط أساليب التقويم مباشرة بنواتج التعليم التى نقومها وبهذا نتيقن من أننا نقوم التلميذ نحو الأهداف التى اخذناها كاطار لعملية التدريس ٠ وعملية ربط اساليب التقويم بنواتج التعلم النوعية هى فى اساسها عملية تحليل منطقى وحكم ٠ هذه العملية يمكن تيسيرها باستخدام خطة ما من خطط التقويم المنظمة ، هذه الخطة يجب أن تشمل على الأقل قائمة بنواتج التعلم المرغوب فيها والأساليب التى تستخدم فى تقويم التقدم نحوها ٠ واليك خطة على سبيل المثال فى ميدان علم الأحياء للصف الأول الثانوى تشـمل الأهداف التعليمية ونواتج التعلم النوعية وأساليب التقويم الملائمة ٠

١ ـ يعرف التلميذ المصطلحات الشائعة المستخدمة فى علم الأحياء حيث:

١ ـ ١ يعرف المصطلحات الشائعة فى علم اساس المعنى ١ ـ ٢ ـ اختبار موضوعى الاجابة القصيرة

١ ـ ٢ يميز بين المصطلحات الشائعة على أساس المبنى ١ ـ ٢ اختبار موضوعى

١ ـ ٣ يتعرف على معانى المصطلحات الشائعة حين تستخدم فى نص معين ١ ـ اختبار موضوعى ٠

٢ ـ يظهر التلميذ القدرة على التفكير الناقد حينما :

٢ ـ ١ يميز بين الحقائق والآراء ٢ ـ ١ اختبار موضوعى

٢ ـ ٢ يشتق استنتاجات صحيحة من المعطيات ٢ ـ ٢ اختبار مقال من نوع الاجابة القصيرة ٠

٢ ـ ٣ يتعرف على الافتراضات وراء الاستنتاجات ٢ ـ ٣ اختبار موضعى

٢ ـ ٤ يتعرف على نواحي القصور فى المعطيات ٢ ـ ٤ اختبار موضوعى

٣ ـ يؤدى التلميذ الأعمال الأساسية للتشريح بمهارة حينما :

٣ ـ يضع العينة فى الوضع الصحيح للتشريح ٣ ـ ١ قائمة ملاحظة أو مقياس تقدير

٣ ـ ٢ يقطع بمهارة دون افساد العينة التى يدرسها ٣ ـ ٢ قائة ملاحظة او مقياس تقدير ٠

٣ ـ ٣ يقسم اجزاء البنية فى العينة دون افسادها ٣ ـ ٢ قائمة ملاحظة او مقياس تقدير ٠

٣ ـ ٤ ينتهى من التشريح فى الوقت المحدد ٣ ـ ٤ قائمة ملاحظة او مقياس تقدير ٠

٤ ـ يظهر التلميذ القدرة على تحديد مواضـــع المعلومات البيولوجية حينما :

٤ ـ ١ يستخدم كتالوج المكتبة فى تحديد المراجع ٤ ـ ١ تقرير بحث أو ملاحظة

٤ ـ ٢ يعرف المصادر المشتركة للمعلومات البيولوجية ٤ ـ ٢ اختبار موضوعى ٠

٤ ـ ٣ يستخدم جداول المحتوى والفهرس عند البحث عن المعلومات فى الكتب ٤ ـ ٣ ملاحظة

٤ ـ ٤ يحدد مدى ملاءمة المعلومات لمشكلة محددة ٤ـ٤ تقرير بحثأوملاحظة

٥ . يكون لدى التلميذ اتجاه علمى نحو الظواهر البيولوجية حينما :

٥ ـ ١ يؤجل الحكم حتى تتوافر جميع الحقائق او معظمها ٠

٥ ـ ١ تقرير قصصى ، اختبار موضوعى

٥ ـ ٢ يبحث عن علاقات السبب والأثر فى البيانات البيولوجية

٥ ـ ٢ تقرير قصصى • اختبار موضوعى

٥ ـ ٣ يظهر رغبة فى التفاسير الجديدة للبيانات البيولوجية

٥ ـ ٣ تقرير قصصى • اختبار موضوعى

٥ ـ ٤ يفسر البيانات البيولوجية متحررا من التحيز

٥ ـ ٤ تقرير قصصى • اختبار موضوعى

٥ ـ ٥ يظهر ثقة فى البيانات البيولوجية التى يحصل عليها بالأساليب العلمية

٥ ـ ٥ تقرير قصصى • اختبار موضوعى

من هذا الجدول نجد ما يأتى :

١ ـ نواتج التعلم النوعية كما تصاغ فى صورة سلوك للتلاميذ عديدة ومتنوعة بحيث لا توجد اداة تقويم واحدة تعطى دليلا مباشرا على تحصيل هذه الأهداف التربوية • ولذلك نحتاج الى جانب الاختبارات الموضوعية الى قوائم الملاحظة والتقارير القصصية واختبارات المقال

٢ ـ تظهر اهمية الصياغة الواضحة للاهداف ونواتج التعلم فى انتقاء اداة التقويم • وفى الواقع حينما تصاغ نواتج التعلم فى ضوء سلوك التلميذ فانها لا تدل فقط على ماذا نقوم ولكنها توحى ايضا بكيف نقوم •

مثال : (١ ـ ١) « يحدد المصطلحات الشائعة » يوضح لنا اسلوب التقويم الذي يجب أن يستخدمه • فهى تدل على ان التلميذ يجب أن يعطى التعاريف بنفسه وعلى ذلك فان اختبار المقال من نوع الاجابة القصيرة حيث يعطى للتلاميذ عدة مصطلحات ويطلب منه تعريفها هو الأسلوب الأكثر ملاءمة لتقويم هذا الهدف • أما الاختبار الموضوعى مثـــلا من نوع الاختيار من متعدد حيث يطلب من التلاميذ أن يتعرفوا فقط على التعريف لا تتناسب مع تقويم هذا الناتج التعلمى فى صيغته هذه ، والا فان الصيغة تصبح « يتعرف التلميذ على معنى المصطلحات الشائعة » • وفى هذه الحالة يمكن استخدام الاختبار الموضوعى • ويتوقف هذا على غرضنا من التقويم • فاذا كنا نريد أن نتأكد من ان التلاميذ يعرفون المصطلحات عن طريق تحديدهم وتعريفهم لها فان الاجراء الوحيد المـــلائم هو أن يطلب من التلاميـــذ أن يعرفوا

مصطلحاتهم • والقدرة على التعرف الصحيح لا يمكن اعتبارها دليلا على قدرة التلاميذ على اعطاء التعريف الصحيح ما لم يكن هناك دليل من البحث العلمى يؤكد أن نوعى السلوك مرتبطان ارتباطا عاليا • وحيث انه لا يوجد مثل هذا الدليل بعد فان الاجراء الأحوط هو اختيار اسلوب التقويم الذى يرتبط مباشرة بنتائج التعلم النوعى (الهدف التعليمى) •

تحليل محتوى المادة الدراسية :

أشرنا الى أن الاختبار التحصيلى يرتبط بمادة دراسية معينة تم تدريسها بالفعل ، ومعنى ذلك أنه لا يوجد ما يبرر اعداد اختبارات من هذا القبيل تتناول مواد لم يتم تدريسها أو لم تكن محددة كمحتوى لعملية التعلم• ولذلك لا بد فى اعداد الاختبار التحصيلى من تمثيل كل ما حظى باهتمامنا اثناء عملية التدريس • ومن الطرق الملائمة فى تحديد الأصل الاحصائى المناسب الذى نشتق منه عينة السلوك التى تسمى الاختبار التحصيلى أن نعد قائمة بالموضوعات التى تتضمنها هذه المادة باكبر قدر من التفصيل وفى ابسط صورة ممكنة • ويمكن للباحث فى هذا الصدد أن يستفيد بمنهج البحث المعروف باسم « تحليل المحتوى » Content analysis (٥) •

تحديد الوزن النسبى للإهداف التعليمية ومحتوى المادة الدراسية :

بعد تحليل المحتوى يقوم الباحث بتحديد الوزن النسبى لكل موضوع (أو وحدة) من موضوعاته ، ومن المحكات المستخدمة فى تحديد وزن كل موضوع من موضوعات المحتوى هو الزمن المخصص لتدريس كل منها • الا أنه فى بعض الأحيان لا يكون الوقت المستغرق فى تدريس موضوع معين دلالة على اهميته ، ولذلك فان الأوزان النسبية للموضوعات لا بد أن تعكس اهتمام المعلم به • وبالمثل لا بد من أن يحدد المعلم الأوزان النسبية للاهداف التعليمية لمادته • فاذا كان يؤكد فى تدريسه وصول التلاميذ الى التعميمات وتطبيقها مثلا فيجب أن يعطى هذا الهدف وزنا اكبر •وبعبارة موجزة لابد ان تعطى لموضوعات المحتوى وللاهداف التعليمية اوزان نسبية تعكس الاهتمام الذى تحظى به فى عملية التدريس •

وفى تحديد هذه الأوزان قد يعتمد المعلم على خبرته فى تدريس المادة وتحليله لمحتواها فيضع قائمة بأهم الموضوعات ثم يصنفها حسب الأهداف التى يرى أن يقيسها الاختبار • الا انه فى كثير من الأحيان لا يكتفى الباحث بالاعتماد على خبرته الشخصية وحدها ويلجا الى بعض الخبراء لاستشارتهم فى تحديد هذه الأوزان النسبية •

اعداد جدول المواصفات :

على اساس المعلومات التى يتم تجميعها يتم تحديد مواصفات الاختبار test specifications والتى تتضمن الموضوعات التى يجب أن يشملها الاختبار ونواتج التعلم التى يجب اختبارها (أى الأهـــداف التعليمية) ، والأهمية النسبية (أو الوزن النسبى) للموضوعات والأهداف • وعلى هذا الأساس يتم تحديد عدد الأسئلة التى يتم اعدادها مرتبطة بكل هدف فى كل موضوع • والطريقة الأكثر سهولة فى تحديد مثل هذه المواصفات هى اعداد جدول المواصفات table of specifications كما تناولناه فى الفصـــل الخامس •

تحديد نوع الأسئلة :

اشهر انواع الأسئلة التى تستخدم فى ميدان الاختبـارات التحصيلية أسئلة المقال والأسئلة الموضوعية ، وأسئلة حل المشـــكلات (او الأسئلة التفسيرية) •ونعرض فيما يلى عرضا موجزا لأنواع الأسئلة الشائعة فى التقويم التربوى الحديث •

أسئلة الاختيار من متعدد .

وتعتبر اكثر الأسئلة الموضوعية شيوعا ، وتقيس بكفاءة شديدة النواتج البسيطة للتعلم ، بل قد تصلح لقياس بعض النواتج التعليمية المعقدة•ويتكون السؤال من مشكلة (قد تصاغ فى صورة سؤال مباشر او عبارة ناقصة) تسمى الجذر Stem ، وقائمة من الحلول المقترحة (قد تشتمل كلمات او اعدادا او رموزا او عبارات) وتسمى البدائل الاختيارية alternatives . ويطلب من المفحوص قراءة جذر الســـؤال وقائمة البدائل وانتقاء البديل

الصحيح أو الأفضل . والبديل الصحيح فى كل سؤال يسمى ببساطة الاجابة answer بينما تسمى البدائل الاختيارية المحـــولات (أو المشتتات) distractors ، وهى تسمى كذلك من واقع وظيفتها المقصـــودة . أى تحويل هؤلاء الذين فى شك من الاجابة الصحيحة .

ويستخدم هذا النوع من الأسئلة فى قياس كثير من الأهداف التعليمية وخاصة نواتج التعلم فى ميدان التذكر واكتساب المعـــلومات ، ومن ذلك معرفة المصطلحات ومعرفة الحقائق النوعية ، ومعرفة المبادىء والقوانين ، ومعرفة الطرق والاجراءات ، كما تستخدم فى قياس الفهم مثل القدرة على تطبيق الحقائق والمبادىء ، والقدرة على تفسير عـــلاقات السبب والأثر ، والقدرة على تبرير الطرق والاجراءات .

ويتميز هذا النوع بمرونته الشديدة ، وتحـــرره من نقائص الأنواع الأخرى من الأسئلة ، ومن مشكلة اســـلوب الاســـتجابة response set بل ان صياغة البدائل الاختيارية قد يؤدى استخدامها فى أغراض التشخيص . ومع ذلك فان لأسئلة الاختيار من متعدد مشكلاتها وأهمها أنها تقتصر على المستوى اللفظى للاداء (وهى فى هذا تشترك مع غيرها من أسئلة الورقة والقلم) ، كما أنها لا تصلح لقياس بعض مهارات حل المشكلة فى الرياضيات والعلوم ، والقدرة على تنظيم الأفكار (وهى فى هذا تشترك مع غيرها من أسئلة التعرف على الاجابة) . ثم ان لأسئلة الاختيار ن متعدد نقيصـــة خاصة بها لا تشترك فيها مع غيرها ، وهى صعوبة ايجاد عـــدد كاف من المحولات المحبذة وغير الصحيحة فى وقت واحد . الا أن هـــذه الصعوبة تتناقص مع زيادة الخبرة فى بناء مثل هذه الأسئلة .

أسئلة التفسير :

يتطلب هذا النوع من الأسئلة ما هو اكثر من الحفظ واكتساب المعلومات الجزئية ، أى انه يستخدم فى قياس نواتج التعلم المرتبطة بالعمليات العقلية العليا كالفهم ومهارات التفكير وقدرات حل المشـــكلة . ويتكون السؤال التفسيرى interpretative من سلسلة من الأسئلة الموضوعية تعتمد على مجموعة مشتركة من البيانات الأولية (المعطيات) . وقد تكون هـــذه

المعطيات فى صورة مواد مكتوبة أو جداول او رسوم أو اشكال او خرائط او صور • وقد تتخذ الأسئلة المرتبطة صورا مختلفة ولكنها فى الغالب تاخذ صورة الاختيار من متعدد •

اسئلة المزاوجة :

تتكون اسئلة المزاوجة matching التقليدية من عمـــودين متوازين يحتوى كل منهما على مجموعة من العبارات أو الرموز او الكلمات • وتسمى المفردات فى العمود الذى ننشد له المزاوجة بالمقدمات premises ، وتسمى المفردات التى نختار منها بالاستجابات responses .

وعادة ما تكون اسس مزاوجة الاستجابات للمقدمات واضحة فى ذاتها، وفى بعض الأحيان تتطلب توضيحا فى التعليمات • ويستخدم هذا النوع فى قياس الحقائق التى تستند الى التداعى البسيط مثل الشخصيات والانجازات، التواريخ والأحداث التاريخية ، المصطلحات والتعاريف ، القواعد والأمثلة ، الرموز والمفاهيم ، المؤلفون ومؤلفاتهم ، الكلمات الأجنبية ومقابلاتها بالعربية الآلات واستخداماتها ، النبات وتصنيفاته ، الحيوان وتصنيفه ، الأشـياء واسماؤها ، الأعضاء ووظائفها • وتتميز بأنها تقيس بدقة وسرعة مقدارا كبيرا من الحقائق المرتبطة فى زمن قصير نسبيا ، بالاضافة الى انها سهلة الصنع • ومع ذلك فان لهذا النوع مشكلاته وأهمها انه يقتصر على قياس الحقائق التى تعتمد على التذكر والتداعى البسيط • وكذلك يصـعب على الباحث ايجاد مادة متجانسة لها أهميتها فى المحتوى الدراسى يمكن أن تصاغ منها قائمة المقدمات •

اسئلة الاجابة القصيرة (الاستدعاء) :

يتطلب هذا النوع من الأسئلة الذى يعرف فى بعض المؤلفات باسم اسئلة التكملة ـ ان ينتج المفحوص استجابته ـ وليس مجرد التعرف عليها كما هو الحال فى الأنواع السابقة • ويمكن القول أن هذا النوع قد يتطلب اجابة قصيرة اذا عرضت المشكلة فى صورة سؤال مباشر ، او تكملة اذا عرضت فى صورة عبارة ناقصة • وتوجد أنواع أخرى من هذا النوع منها : أسئلة

اعداد القوائم Listing والتى تسمى احيانا اختبارات المقال ذات التعليمات المحددة (مثل اذكر ٣ اسباب لقيام الحرب العالمية الأولى) واسئلة التماثل او التناسب analogy . واسئلة المشكلات ، واسئلة التعيين identification (كأن يطلب من التلميذ تعيين اجزاء رسم نبات أو حيوان) . ويستخدم هذا النوع فى قياس النواتج البسيطة نسبيا للتعلم مثل معرفة المصطلحات ، ومعرفة الحقائق النوعية ، ومعرفة المبادىء ، ومعرفة الطرق والاجراءات ، والتفسيرات البسيطة ، والقدرة على حل المسائل فى الرياضيات والعلوم . وهذا النوع ربما بسبب بساطة نتائج التعلم التى تقيسها (الاستدعاء فى اغلب الأحوال) ، بالاضافة الى تحرره من اثر التخمين السائد فى اختبارات التعرف بوجه عام . الا أن اكثر مشكلاته شيوعا هى صعوبة تصحيحه .

اسئلة الصواب والخطأ (او اسئلة البديلين alternative response) :

وتتطلب اختيار اجابة واحدة من اجابتين كالحكم على العبارة بالصواب أو الخطأ ، أو الاجابة على السؤال بنعم او لا ، أو الحكم على العبارة بأنها تدل على رأى أو حقيقة ، و تقدير عبارة بالموافقة و المعارضة . ويستخدم هذا النوع فى قياس نتائج التعلم التمييزى البسيط ، ومع ذلك فيمكن أن تفيد فى قياس الفهم (كالتعرف على علاقات السبب والأثر) . وقد يكون اكثر مشكلات هذا النوع من الأسئلة خطورة هو تأثرها البالغ بالتخمين .

اسئلة الترتيب :

وفيها يقوم المفحوص باعادة ترتيب خطوات او احداث او اجراءات او تواريخ فى تسلسل طبيعى أو منطقى .

اسئلة المقال :

رغم الاستخدام الواسع للاسئلة الموضوعية التى اشرنا اليها لازالت توجد نتائج هامة للتعلم لا يصلح لها الا سؤال ، المقال ، ومنها القدرة على عرض وتنظيم وتكامل الأفكار ، والقدرة على التعبير الكتابى ،

والقدرة على اعطاء التفسيرات والتطبيقات للمعلومات لا مجرد استدعاؤها او التعرف عليها ·

ومن اهم خصائص اسئلة المقال حرية الاستجابة ، وفى هذا جداوها كمقياس للتحصيل المعقد · وفيه أيضا تكمن صعوبات التصحيح التى تجعل منها اداة اقل كفاية وكفاءة فى قياس الحقائق والمعلومات · ومعنى ذلك أن اسئلة المقال يجب أن تستخدم حيث لا تصلح الاختبارات الموضوعية ·

الاختبارات العملية :

تستخدم هذه الاختبارات كوسائل موضوعية لتقدير الكفاءة التى يؤدى بها العمل (كالكتابة على الآلة الكاتبة) أو قيادة السيارات ·وتنقسم هذه الاختبارات الى ٣ انواع :

(ا) اختبارات التعرف : وتتطلب من المفحوص التعرف على الخصائص الأساسية للاداء (كأن تعزف مقطوعة موسيقية ويطلب من المفحوص تحديد الأخطاء أو النقائص فى الأداء أو التعبير) ، أو تحديد الأجزاء التى تتألف منها احدى الآلات ، أو اختيار الآلة أو الجهاز المناسب لعمل معين ، أو تحديد العينات او تصنيف الأشياء ، أو انتقاء عمل أدبى أو فنى ممتاز ·

(ب) الاختبارات التى تتضمن مواقف تشبه المواقف الطبيعية : فهى تهدف الى قياس الأنشطة الأساسية فى العمل · وتسمى أحيانا اختبارات النماذج المصغرة miniature ·

(حـ) اختبار عينة العمل : وهو عبارة عن محاولة « مضبوطة » أو « مقننة » فى الظروف الواقعية للعمل ، وتنقسم هذه الاختبارات الى نوعين اساسين : اولهما الاختبارات التى يسهل فيها التمييز بين الصواب والخطأ فى الأداء وبالتالى يمكن تصحيحها بسهولة مثل التصويب ، والأداء العضلى فى التربية البدنية ، والتجميع الميكانيكى ، والكتـابة على الآلة الكاتبة · والنوع الثانى هو الاختبارات التى تعتمد على حكم المراقبين والفاحصين لتقويم الأداء واعطاء درجة أو رتبة ، ومن امثلة ذلك محاولة سامى ابراهيم (التقويم النفسى)

على (٤) ، تقنين ، اختبارات العزف على البيانو • ويتطلب هذا النوع استخدام مقاييس التقدير أو قوائم الملاحظة •

الاختبـــارات الشفوية :

الاختبـــار الشـــفوى هو مزيج من اختبـــار المقـــال والاختبـــار العملى ، وله فائدته فى دراســـة العمليـــات العقليـــة التى يستخدمهـــا المفحوص فى الاجابة على أسئلة معينة ، ولذلك فهو أداة نافعة فى تشخيص الصعوبات • بل انه فى بعض الأحوال هو الأسلوب الأوحـــد فى تقويم المفحوص كما هو الحال فى قياس قدرات صغار التلاميذ (فى الصفوف الأولى من المرحلة الابتدائية) ، وفى قياس بعض نتائج التعلم اللغوى (كالقراءة الجهرية) •

تحديد عدد الأسئلة :

يتحدد عدد الأسئلة التى يتكون منها الاختبار بمقدار الزمن الميسر له • وكثير من الاختبارات التحصيلية يتحدد زمنه بحوالى ٥٠ دقيقة (زمن الحصة المدرسية العادية) وبعض جداول الامتحانات تحدد فترات أطول تصل الى ٣ ساعات • وعلى وجه العموم كلما طال الزمن وزاد عـــدد الأسئلة كان الاختبار اكثر ثباتا • ومن النادر من الوجهة العملية أن نعد اختبارا ليستغرق اكثر من ثلاث ساعات •

ومن المفيد كما أشرنا أن نعتبر مجموعة الأسئلة التى يكون منها الاختبار عينة من أصل افتراضى لكل الأسئلة الممـــكنة التى يمكن أن تســـتخدم فى الاختبار • فمدرس الصف الخامس الابتدائى قد يحصـــل على ١٠٠ كلمة لاختبار المفردات اللغوية وذلك بانتقاء الكلمة الخامسة من قائمة تحتوى على ٥٠٠ كلمة درست أثناء العام الدراسى • وفى هذه الحالة تصبح الـ ٥٠٠ كلمة هى الأصل الذى تؤخذ منه عينة الاختبار ، وهو أصل حقيقى ومحدد المعالم • الا أنه فى معظم الاختبارات لا نجد مثل هذا التحديد الحاســـم للأصل الاحصائى العام للاسئلة ، فمثلا لا يوجد حد أقصى لعدد المشـــكلات والمسائل المختلفة التى يمكن اعدادها فى اختبارات الجبر •

وبالطبع كلما اتسع مدى الموضوع أو القدرات التى يهـــدف الاختبار

الى قياسها زاد عدد الأصل الاحصائي · وحجم هذا الأصل يمثل الحد الأقصى لحجم عينة الاختبار · ومعنى ذلك أن أسئلة الاختبار لا يمكن أن تكون أكثر عددا من الأصل التى تشتق منه ، الا أن حجم هذا الأصل لا يصنع حدا أدنى لحجم العينة · فالأصل الذى يتألف من ١٠٠٠ سؤال يمكن أن تكون العينة المشتقة منه ١٠ أسئلة أو ٥٠ سؤالا أو ٢٠٠ سؤال وكلما زاد حجم الأصل قل تجانسه ، أى أصبح يحتوى ميادين مختلفة وشبه مستقلة من المعرفة والقدرة · وللوصول الى نتائج دقيقة يجب أن تكون عينة الاختبار المشتقة من أصل غير متجانس أكبر نسبيا من تلك التى تشتق من أصل متجانس · ومع ذلك فلا توجد قاعدة عامة تحدد العلاقة من حجم الأصل والعينة ·

ويوجد فى الوقت الحاضر اتجاه يؤكد استخدام عدد قليل وكاف من الأسئلة بحيث يعطى معظم المفحوصين وقتا يحاولون فيه جميع الأسئلة بمعدلاتهم الخاصة العادية · ولشيوع هذا الاتجاه ثلاثة أسباب (١٣) ·

١ - ان سرعة الاستجابة ليست هدفا أوليا للتعلم فى معظم المقررات الدراسية وبالتالى لا تعد مؤشرا صادقا للتحصيل لأنه فى كثير من ميادين التحصيل لا ترتبط السرعة والدقة ارتباطا عاليا · وبالطبع فقد تكون اختبارات السرعة أكثر ملاءمة فى بعض الحالات ، الا أن هذه المواقف هى الاستثناء وليس القاعدة ·

٢ - ان قلق الاختبار قد يكون على درجة كبيرة من الحدة - حتى فى الاختبارات غير الموقوتة - حين يتزايد التأكيد على العمل بسرعة ·

٣ - ان الاستخدام الكفء للاختبار التحصيلى - الذى يصنعه المعلم - يتطلب أن معظم التلاميذ يستجيبون له كله حتى يستخدم نتائجه فى أغراض التشخيص ·

تقويم التحصيل فى القراءة فى المدرسة الابتدائية (*) :

قد تكون القدرة على القراءة أهم القدرات التى تؤكدها المدرسة

(*) الامثلة التى سوف نعرض لها مأخوذة بكثير من التعديل عن لندفال (١٧) ·

الابتدائية · وفى الصفوف الأولى من المرحلة الابتدائية يحتل تعليم القراءة كمهارة أساسية الجزء الأكبر من العمل المدرسى ، ومع انتقال التلميذ الى الصفوف العليا تتفرع دراسته الى مجالات متنوعة ويقل الوقت المخصص لتعليم القراءة ، ومع ذلك فان لها اهميتها طوال المرحلة الابتدائية · وبعدما يتقن التلميذ بعض المهارات الأساسية فى القراءة ، قد تتعدل تعديلات طفيفة أهداف تدريس القــــراءة بحيث تؤكد التغيرات الناتجة فى المهارات التى نقيسها · ومن ناحية أخرى فانه لا زال من الأهداف الأساسية للتعلم فهم معانى الكلمات وفهم الجمل والفقرات الطويلة ، وهى جميعا لها اهميتها فى جميع المستويات ، ومعنى هذا أن اى اقتراحات تتعلق بتقويم أهداف القراءة فى أحد صفوف المرحلة ينطبق على كل المرحلة ·

وفيما يلى مثال لأهداف نوعية لتعليم القراءة فى الصف الأول الابتدائى

الغرض العام	الهدف المرحلى (فى نصف العام)	الأهداف النوعية
أولا : أن يســتطيع الفرد المتعلم قراءة اللغة القومية ، قراءة فعالة ·	١ ـ أن يستطيع التلميذ قراءة القســم الأول من كتــاب القراءة بسرعة وفهم	١ ـ أن يستطيع الطفل أن ينطق بكلمات مثل مدرســة ، الحديقة ، ورد ، حينمــا يراها مكتوبة (عــــدد كلمات القســم الأول من القراءة ٢٥٠ م) ·
		٢ ـ أن يستطيع الطفل أن يربط كلمات مثل ولد ، طيارة ، بيت وبين الصــور التى تعبر عنها ·
		٣ ـ أن يستطيع الطفل أن يجيب على أسئلة تدور حول مواد جديدة يقرأها من مختارات من كتاب القراءة ·
	ب ـ أن يستطيع الطفل القــراءة الجهرية بســهولة مع بعض التعبير	٤ ـ أن يســتطيع الطفل أن يغير من نبرات الصــوت للتمييز بين الجملة الخبرية والجملة الاستفهامية ·

ومن الواضح انه لو كان الواجب هو تقويم هذه الأهداف النوعية جميعا فيجب أن تستخدم اساليب مختلفة • فالهدفان ١ ، ٤ يمكن تقويمها ببعض صور الأداء الشفوى • بينما الهدفان ٢ ، ٣ يمكن تقويمهما بهذه الطريقة أو باستخدام بعض المواد المكتوبة • وبسبب الفوائد الكثيرة الخاصة باستخدام الاختبارات الموضوعية فى المدرسة الابتدائية فان على المعلم أن يختار لقياس تحصيل الهدف النوعى رقم ٢ طريقة المزاوجة بين صور الأشياء والكلمات التى تدل عليها • وفى هذا المستوى يجب أن يعطى لجميع التلاميذ تعليمات شفوية تقرر انهم يجب عليهم أن يرسموا خطا يصل بين الكلمة والصورة التى تدل على معنى الكلمة • ومثل هذا السؤال يعطى قياسا مباشرا لقدرة الطفل على ربط كل كلمة بالصورة الملائمة • والعلاقة البسيطة بين الهدف وهذا الاختبار تدل على فائدة صياغة الأهداف فى عبارات تدل على سلوك يمكن ملاحظته •

وبالنسبة لتقويم الهدف ٣ فان المعلم يطلب من تلاميـــذه الانتقال الى مواد جديدة وقراءتها قراءة صامتة • وبعد هذه القراءة توجه اليهم اسئلة يطلب منهم كتابة اجابات عليها وتتعلق هذه الأسئلة بمحتوى للفقرة • وهذا الاختبار قد يأخذ صورة أن يطلب من التلاميذ أن يكتبوا جملة فى الاجابة على السؤال - وهو صورة مبسطة لاختبار المقال ، الا أن التلاميذ فى هـــذا المستوى لا نتوقع لهم أن تكون عندهم سهولة كبيرة فى الكتابة ، ولذلك فان بعض الاختبارات الموضوعية ينتج قياسا اكثر صدقا لفهم القراءة عندهم مثل اسئلة الاختبار من متعدد مثل •

فى يوم العيد ذهبنا الى : الحقل

المدرسة

الحديقة

ومرة أخرى يجب أن تكون تعليمات هذه الأسئلة شفوية وتطلب من التلميذ ان يضع خطا أو يرسم دائرة حول الاجابة الصحيحة للسؤال •

وحيث ان القراءة هى مهارة اتصال ، فان جميع أهداف التعليم فى هذا الموضوع تهتم اساسا بالفهم • وخاصة فهم معانى الكلمات • فمن المتوقع من التلميذ ان يشرح معنى كلمة ، أو يختار مرادفا لها ، أو يربط بينهـا وبين

الصورة أو الشيء الذى تصفه • كما تربط الأهداف الأخرى بينهم الجمل او الفقرات •

ومع تقدم الطفل فى صفوف المرحلة الابتدائية فان التأكيد على التحصيل سى القراءة يتزايد اهتمامه بانماط مختلفة وجوانب مختلفة من الفهم ، فتتطلب اختبارات الفهم من التلميذ اكثر من مجرد استدعاء اجـــزاء من معلومات الحقائق والتفاصيل •

تقويم التحصيل فى الحساب :

تتزايد اهمية المهارات الحسابية والفهم الحسابى فى الحضارة المعاصرة تزايدا مستمرا • ومن المهم لتلاميذ المرحلة الابتدائية أن يتلقوا تعليما جيدا فى المهارات الأساسية فى الحساب وفى عناصر التفكير الكمى • وللمساعدة فى تحقيق هـــذا الهدف من المهم للمعلم أن يقوم بتقويم مستمر لتقدم كل تلميذ • وهذا مهم على وجه الخصوص فى الحساب حيث كل مهارة لاحقة تعتمد على المهارة السابقة ، وحيث يفشل التلميذ اذا فشل فى اتقان أى من المعلومات الأساسية مثل :

الغرض العام	الهدف المرحلى	الهدف النوعى
أولا : أن يســـتطيع الشخص المتعلم حـل مشـــكلات العـــد والحساب	أ ـ أن يستطيع التلميذ أن يعـــد حتى الرقم ١٠٠ •	١ ـ أن يستطيع التلميذ أن يعد ويسجل العـــدد الصحيح لأشـــياء أو حروف توضـحها الصور فى حدود ١٠ وحدات ٢ ـ أن يستطيع الطفل العد الى الرقم ١٠٠ بزيادة واحد•
	ب ـ أن يعرف التلميذ حقائق الجمع الأساسية	٣ ـ أن يستطيع التلميـــذ الاجابة على مســـائل الجمع لأزواج من الأعداد التى تتكون من رقم واحـــد بحيث يكون مجموعها أقل من ١٠

د - أن يستطيع التلميذ فهم المصطلحات الحسابية الخاصة	٤ - أن يستطيع التلميذ شرح معنى مصطلحات مثل جمع وطرح ومجموع وأكثر وأقل ·
	٥ - أن يستطيع التلميذ أن يختار اختيارا صحيحا للأشياء التي تتحدد بمصطلحات مثل الأول والثاني والثالث··الخ·

وجميع هذه الأهداف النوعية الخمسة يمكن تقويمها اما شفويا أو تحريريا ، فالهدف الأول يمكن قياسه باستخدام أسئلة بسيطة يعطى فيها للطفل أعدادا مختلفة من الأشياء عليه أن يعدها ·

تعليمات (شفوية) : اكتب على الخط الموضوع أسفل كل صورة العدد الذى يدل على عدد الحروف فى كل منها ·

د د	ص ص	ح ح	ه ه ه
د	ص ص ص ص	ح ح	ه ه
	ص ص	ح ح	ه ه ه
ــــ	ــــ	ــــ	ــــ

ويمكن تقويم الهدف الثانى بسؤال التلميذ أن يعد جهريا أو أن يكتب الأعداد من ١ الى ١٠٠ ·

والهدف الثالث يمكن قياسه بكفاءة أكثر بمجرد سؤال التلميذ أن يكتب اجاباته على عدد من المسائل التى تتضمن الجمع البسيط :

٣	٧	٥	١	٦
+ ٤	+ ٢	+ ٤	+ ٨	+ ٣
ــــ	ــــ	ــــ	ــــ	ــــ

والهدف الرابع يمكن تقويمه بسؤال التلميذ مباشرة أن يشرح هذه

المصطلحات شفويا أو تحريريا · ويوجد تعديل طفيف لهذا الهدف بحيث يتطلب من التلميذ أن يتعرف على صور أشياء أو عمليات يمكن أن توصف لفظيا كأن يعطى له رسم لعدد من الكرات متفاوته الحجم ويطلب منه وضع خط تحت الكرة الأصغر ·

والهدف الخامس يمكن تقويمه باستخدام اسئلة مماثلة ، كأن يعطى له رسم لعدد من الكرات المتساوية ويطلبمنه وضع خط تحتالكرة الرابعةمثلا·

ان أنواع الأسئلة التى ذكرناها حتى الآن تتعلق أساسا بقياس معلومات التلميذ وتذكره لحقائق خاصة أو مصطلحات نوعية · الا أن المعلم قد يهتم بأنماط أخرى من القدرات وخاصة تلك التى ذكرناها فى تصنيف الأهداف التربوية (لبلوم) وتحت عناوين مثل الفهم والتطبيق ·

مثال لترجمة الأعداد والمفاهيم احداها الى الآخر :

١ ـ ما الذى يعطى نفس الاجابة التى يعطيها ٥ × ٤ ؟

أ) ٤ + ٤ + ٤ + ٤ + ٤
٥

(د) ·

(أ) $\frac{٦}{٨}$ (ب) $\frac{٥}{٨}$ (حـ) $\frac{٤}{٨}$ (د) $\frac{٣}{٨}$

ومن أهم الأهداف التعليمية التى تحتـــاج الى تقويمها فى اختبارات الحساب هى القدرة على تطبيق معرفة المبادىء والأساليب أو الاجراءات (وتقاس باختبارات المسائل الحسابية من مختلف الأنواع) ·

تقويم التحصيل فى العلوم :

مع زيادة التأكيد على العلم فى منهج المدرسة الابتدائية ككل ، نجد من

الطبيعى أن تسعى المدرسة الابتدائية الى زيادة مدى ما يتضمنه برنامجها من العلوم. والأهداف التى نسعى الى تحقيقها يجب أن تكون اكثر من مجرد معرفة حقائق معينة أو مصطلحات أو حفظ مبادىء وقوانين . وانما يجب أن تتضمن قدرات مثل القدرة على اجراء تجارب بسيطة ، واشتقاق المبادىء والتعميمات من البيانات أو اعطاء أمثلة للمبادىء المختلفة وتطبيق مثل هذه المبادىء فى شرح النتائج أو التنبؤ بها كما يجب أن تتعلق الأهداف باتجاهات وميول معينة .

وحيث أن أغلب معلمى المدرسة الابتدائية يعوزهم الاعداد السليم فى العلوم فانهم يحتاجون الى المساعدة فى تحديد أنماط التعلم التى تنتج عن مقرر العلوم تحديدا واضحا وفى الحصول على أمثلة خاصة من القدرات المختلفة . ومن الأساليب المفيدة فى هذا الصدد اعداد وانتقاء اختبارات جيدة وأسئلة جيدة فى الاختبارات ، فكثير من القدرات يمكن التدليل عليها بطريقة جيدة عن طريق تطبيقها فى مواقف تعد أسئلة الاختبار أمثلة لها أو مواقف تتحول الى أسئلة اختبار . ومعرفة الطريقة التى يمكن بها تقويم هذه القدرات يساعد فى توضيح المعنى الحقيقى للقدرة عند العلم وتشجعه على تدريسها .

وفيما يلى مثال :

الغرض العام	الهدف المرحلى	الأهداف النوعية
أولا : على المواطن المتعلم أن يقدر مصادر الأمة .	(أ) أن يفهم التلميذ أهمية المصادر الطبيعية مثل الماء والأرض والنبات والحيوانات	١ - أن يستطيع التلميذ أن يذكر أسباب ضرورة المحافظة على مصادرنا من الماء والأرض والنبات والحيوان . ٢ - أن يستطيع التلميذ تحديد الطرق المختلفة للمحافظة على المصادر الطبيعية .
ثانيا : يقيس المواطن المتعلم التقدم العلمى باسهامه فى الرفاهية العامة .	(أ) أن يشعر التلميذ ببعض الاسهامات الهامة التى يقدمها العلم فى مجال الرفاهية العامة	١ - أن يستطيع التلميذ أن يتعرف على اسهامات علماء مثل باستير واديسون وسولك ومركونى .

ثالثا : تتوفر لدى المواطن المتعلم وسائل دفاع ضد الدعاية .	(أ) أن يفهم التلميذ الطريقة العلمية فى حل المشكلات .	١ ـ أن يستطيع التلميذ أن يحل مشكلات العلم بالوصول الى النتائج من الأدلة والشواهد .
	(ب) أن يصل التلميذ الى قرارات على أساس الأدلة والشواهد سواء فى مسائل العلم أو المسائل غير المتعلقة بالعلوم .	٢ ـ عندما يطلب من التلميذ ان يتخذ قرارات فى مواقف فرضية فى (مجالات العلوم وغيرها) فانه يصل الى تلك التى تدعمها الشواهد والأدلة .

ان الهدفين النوعيين المشتقين من الغرض العام الأول يهتمان بمعرفة المبادىء والطرق . ويمكن تقويم الهدف الأول بمجرد سؤال التلميذ ان يذكر قائمة ـ اما شفويا أو كتابة ـ بعدد معين من الأسباب التى تتعلق بالمحافظة على كل مصدر من المصادر الطبيعية المذكورة .

ويمكن للهدف الثانى أن يصاغ على نحو يشبه الهدف الأول ، وبالتالى يمكن تقويمه بأن تطلب من التلميذ ان يذكر قائمة شفوية . ويمكن ايضا تقويم هذا الهدف باستخدام الاختبارات الموضوعية .

مثال : حتى يستطيع الفلاح ان يحصل من أرضه على محصول وفير عليه أن يزرع فى حقله .

(أ) نفس المحصول عاما بعد عام .

(ب) محصولا واحدا لثلاث سنوات متوالية ثم محصولا آخر للسنوات الثلاثة التالية :

(جـ) محصولا واحدا لخمس سنوات ثم يليه محصول آخر للسنوات الخمس التالية .

ويمكن صياغة هدف رقم ١ بحيث يسمح للتلاميذ بانتقاء الأسباب

الهامة للمحافظة على المصادر الطبيعية ، وعلى هـــذا النحو يمكن تقويمه باستخدام أسئلة الاختيار من متعدد • أما الهـــدف النوعى المعتمد عــلى الغرض العام الثانى فيمكن اختباره باستخدام أسئلة بسيطة مثل أســئلة المزاوجة بين أنواع الاسهامات التى قام بها مختلف العلماء •

والأهداف المشتقة من الغرض العام الثالث أصعب فى تقويمها نوعا ما • ويمكن أن تفيد فى ذلك الاختبارات الشفوية واختبارات المقال ، حيث يعرض على التلميذ مواقف مشكلة معينة ثم يطلب منه اعطاء الاستنتاجات والحلول المناسبة • كما يمكن تقويم هذه الأهداف باستخدام أسئلة الاختيار من متعدد من النوع التفسيرى كما هو مبين فى المثال الآتى :

مثال : قام أربعة تلاميذ بالاشتراك فى تجربة لاستنبات الفول كما يلى:
قام المعتز بمسح بعض حبوب الفول بقطعة من القطن ثم تركها جافة •
قامت مها بمسح بعض حبوب الفول بقطعة من القطن ثم تركتها مبللة •
قام خالد بوضع بعض حبوب الفول فى طين جاف وتركها جافة •
قام أحمد بوضع بعض حبوب الفول فى طين مبلل وتركه مبللا •

ومن نتائج هذه التجربة حاول تلاميذ الفصل الحكم على أى النتائج الآتية حول الصحيح :

(أ) تنبت حبوب الفول اذا تركت فوق قطعة من القطن الجاف •
(ب) تنبت الحبوب اذا تركت فوق قطعة من القطن المبلل •
(حـ) تنبت الحبوب اذا تركت فى الطين الجاف •
(د) تنبت الحبوب اذا تركت فى الطين المبلل •
(هـ) تحتاج الحبوب الى أن تظل مبللة حتى تنبت •
(و) تحتاج الحبوب الى أن تظل جافة حتى تنبت •
(ز) تحتاج الحبوب الى التربة والماء حتى تنبت •

وبالنسبة لكل من العبارات الآتية افترض انها تصف النتائج الحقيقية للتجربة وعليك أن تقرر أى الاستنتاجات السابقة هو الأقرب الى الصواب

اذا كانت النتيجة واحدة من هذه العبارات ٠ حدد اختيارك بكتابة الحروف أ ، ب ، حـ ، د ، هـ ، و ، ز على الخط القصير أمام العبارات الخاصة بالنتيجة:

١ - حبوب المعتز ومها انبتت ، ولم تنبت حبوب الآخرين ٠٠٠٠

٢ - حبوب أحمد انبتت ، ولم تنبت حبوب الآخرين ٠٠٠٠

٣ - حبوب مها أنبتت ، ولم تنبت حبوب الآخرين ٠٠٠٠

٤ - حبوب مها وأحمد انبتت ، ولم تنبت حبوب الآخرين ٠٠٠٠

٥ - حبوب خالد وأحمد انبتت ولم تنبت حبوب الآخرين ٠٠٠٠

٦ - حبوب المعتز وخالد أنبتت ، ولم تنبت حبوب الآخرين ٠٠٠٠

وبهذه الطريقة يمكن قياس بعض جوانب القدرة على التفكير العلمى فى طفل المرحلة الابتدائية ٠

الفصل الثاني عشر

الاختبارات المرجعة الى المحك

مقدمة :

يغلب على الاختبارات المنشورة أنها اختبارات معيار ، وأسباب غلبة هذا النوع من الاختبارات المنشورة المرجعة الى المعيار أنها مصممة ومعدة وموضوعة بواسطة آخرين غير مستعمليها ٠ وانما مبنية بحيث تكون قيمتها الكبرى فى كيف يتباين أداء جماعات مختلفة عليها٠

ولكن من الممكن ، مع هذا ، أن تفسر درجات بعض الاختبارات فى ضوء عدد من العناصر التى يجيب عليها التلاميذ اجابة صحيحة بصرف النظر عن مقارنة هذا العدد بأداء الجماعة ٠ وفى مثل هذه الحالة نقول ان الاختبار يمثل محكا أدائيا performance criterion، ونقول ان التلاميذ الذين يجيبون على كل عنصر من عناصر الاختبار اجابة صحيحة ، أو يجيبون اجابة صحيحة من عدد محدد سلفا من العناصر فى الاختبار هؤلاء التلاميذ يمكن ان نقول انهم استوفوا حق الأداء الكامل المطلوب من الاختبار ٠ هذا الاتجاه هو الذى يسمى الترجيع الى المحك criterion-referenced أو الترجيع المحكى criterion-referencing .

ويجب ملاحظة أننا لا نستطيع أن نسمى أى اختبار اختبارا مرجعا الى المحك لمجرد أنه ليس اختبارا مرجعا الى المعيار ٠ ذلك لأن معنى الترجيع الى المحك انه الأداء على الاختبار مربوط أو متصل بمراجع السلوك (*) behavior referents ، وأن يكون الاختبار مصمما ومنشأ على هذا الاساس (*) ٠ أى أن الاختبار بحكم طريقة تصميمه ، يجب أن يمدنا بمعلومات

(*) النطاق : الاداءات التى يتطلبها المرمى ، والتى منها نختار بنود الاختبار ٠

عن قدرة التلميذ على القيام بأداءات معينة بصورة مطلقة in absolute terms فاذا كان الاختبار اختبار جمع وطرح الكسور مثلا ، فيجب أن يكون لدينا أساس ما لزعمنا أن التلميذ يعرف كيف يجمع ويطرح الكسور ، اذا استطاع أن يجيب عن عناصر الاختبار كلها ، أما اذا لم يستطع أى تلميـذ أن يتم الاجابة ، واذا كان للاختبار حقا محك أدائى سلوكى ، فانه من الممكن أن نخلص الى انه ما من تلميذ استطاع أن يحقق محك الكفاءة proficiency criterion للمهارة المقاسة ٠

مصادر المحـك :

ولكن من أين يأتى المحك الذى يرجع اليه الاختبار ؟

يقترح كوكس وفارجاس (١٩٦٦) Cox & Vargos أن محكا رئيسيا للترجيع هو أن التدريب يؤدى الى زيادة الكفاءة ٠ فاذا أعطى اختبار لجماعة ما قبل التدريب وبعده ، وكانت نتائج الجماعة بعد التدريب افضل منها قبله فان هذا يعنى أن الاختبار حساس لنواتج التدريب ٠ أى انه من الممكن تفسير نتائج الاختبار على أنها تدل على اكتساب المهارات والمعرفة التى قصد التدريب الى انتاجها ٠ مثل هذا الاختبار يكون اختبارا مرجعا الى المحك ٠

وانه من المؤسف حقا أن يكون عدد قليل جدا من الاختبـارات التى ينشئها المدرسون اختبارات مرجعة الى المحك يمكن أن تتصف بالقوة والمتانة٠ ذلك انه حتى يمكن اعتبار اختبار ما مرجعا الى المحك ، من الضرورى أن نحاول أن نحدد العلاقة بين الأداء على هـذا الاختبار وسـلوك محكى criterion behavior ٠ وحتى يكون لدينا تبرير سليم لاعتبار اختبار ما اختبارا مرجعا الى المحك ، فاننا لابد أن نتبع الخطـوات التالية فى بنائه وترجيعه :

١ ـ يعد تخطيط محتوى content outline يذكر فيه المهـارات والمعرفة التى يعتبر الاختبار محاولة لقياسها ٠

٢ - يميز الأداءات ، اى المرامى القابلة للقياس measurable objectives التى سوف يكون المفحوص قادرا عليها · مع التسليم بان هذا المفحوص قد اكتسب كفاءة فى المهارات والمعرفة المقاسة بالاختبار ·

٣ - يميز المجال أو النطاق domain الذى يحدده كل مرمى ، ويكتب العناصر أو البنود فى الاختبار وفق التفاصيل المعينة فى ذلك المجال أو النطاق ، ثم يختار عشوائيا عنصرين لكل مرمى ليبقى منها الاختبار ·

٤ - يتأكد من صحة حقيقة أن المهارات والمعرفة المقاسة بالاختبار هى فى الواقع لازمة prequisite للمرامى الأدائية أو المرامى السلوكية المميزة المطلوبة فى الخطوة رقم ٢ · وقد يكون هذا هو اكثر جوانب هذه العملية تغيرا ، ذلك لأن عملية اثبات صدق الاختبار validation تبدأ بحكم الشخص الذاتى ، أى بإجراء صدق ظاهرى ، ثم تمتد لتتضمن حكم مجموعة من الخبراء أو تتضمن بيانات أو معطيات حقيقية حصلنا عليها من اعطاء الاختبار لمجموعة اثبتت اداءها الكفء لترى مـا اذا كانت لدى افرادها المهارات والمعرفة الخاصة بالاختبار ·

٥ - يقرر أو يحدد درجة قطع (أو درجة قاطعة) cutoff score أو درجة محك criterion score تعبر عن الأداء على الاختبار الذى يجب أن يحصل عليه المختبر ليظهر انه قد أحرز كفاءة كافية فى المهارات والمعرفة ليكون قادرا على أداء الوان السلوك المحكية criterion behaviors

ملامح الاختبارات مرجعية المحك

وعلى هذا فان اهم ملامح الاختبارات المرجعة الى المحك هى :

١ - انها مؤسسة على عدد من المرامى الأدائية أو السلوكية والتى تعتبر هذه الاختبارات محاولة لقياسها ·

٢ - وانها مصممة بحيث يكون لها درجة عالية من الملاءمة appropriateness لأنها قائمة على مرام · ومحتواها يرتبط بمراميها ·

٣ ـ أنها تمثل عينات من السلوك أو الأداء الواقعي الحقيقي .
٤ ـ أن الأداء عليها يمكن أن يفسر في ضوء درجات قاطعة محـــددة ســـلفا .

وانه من الأهمية بمكان أن نوضح أنه ليست هناك فواصل قاطعة بين الاختبارات المرجعة الى المعيار والاختبارات المرجعة الى المحك . وفي الواقع من الممكن أن نفسر اختبارا ما متجهين أحد الاتجاهين أو كليهما على شرط ان تكون المرامي أو تصنيفات المحتوى المتميزة هي الأساس الذي استعمل في كتابة عناصر الاختبار . على أن الفرق الأساسي بين الاختبارات المرجعة الى المحك والاختبارات المرجعة الى المعيار هو أن الأولى مؤسسة على درجـــة قاطعة محددة سلفا، مفترضا أنها داخلة في صميم الأداء نفسه intrinsic أو هي وثيقة الاتصال بذلك الأداء ، بينما النوع الأخير من الاختبارات مؤسس على أداء جماعة معيار ، أي أن تفسير درجة الاختبار اســـــاسه خارجي extrinsic .

انشـــاؤها واستعمالاتها

والترجيع الى المحك أكثر مناسبة في تطبيقة في مجال اختبار المهارة ، وأشد صعوبة في تطبيقه لقياس السلوك المعقد مثل التفكير وحل المشكلة ، كما يذهب (Ebel, 1970a, Cronbach, 1977) . وعلى الرغم من هذا فان ثمة اهتماما وشغفا باستخدام الترجيع الى المحك استخداما أوســـع في قياس التحصيل المدرسي ، سواء بالنسبة الى الاختبارات التي ينشئها المدرس ، أو الاختبارات المنشورة .

ويستخدم الترجيع الى المحك عن طريق :

١ ـ وضع ، واختبار ، مجموعة من المرامي تمثل النواتج المرغوب فيها للتدريس .

٢ ـ تصميم عناصر أو بنود ، أو اختبارها بحيث نقيس كل مرمى قياسا دقيقا، أي تكون هذه العناصر أو البنود ممثلة لنطاق المرمى objective domain

٣ _ التحديد المسبق presetting لمستويات أداء مقبولة ٠

٤ _ اعطاء الاختبار لمجموعة من التلاميذ وتقويم أدائهم فى ضوء عدد من المرامى التى يستطيعون أداء ما تتطلبه من أداء بنجاح ٠

ويتضمن الترجيع الى المحك مهمة كتابة أو وضع مرام وعناصر اختبار لقياس هذه المرامى ٠ وبينما نجد هذا يسيرا على المدرسين فى فصولهم ، فان كتابة مرام وعناصر لتستخدم على نطاق أوسع ، لمنطقة أو محافظة مثلا ، امر على درجة كبيرة من التعقيد ، وربما كانت مؤسسة اختبارية اكثر صلاحية وقدرة على القيام بها من المدرس ٠ الا انه حتى المؤسسة الاختبارية ذاتها سوف تجد من الشاق أن تعد اختبارا مرجعا الى المحك على أساس قومى أو حتى اقليمى ، ذلك لأن مثل هذا الاتجاه سوف يضحى بتصويب الاختبار نحو حاجات نظام مدرسى بعينه ٠ واذا كانت الحقيقة هى أن المدخلات المحلية اكثر ضرورة للاستخدام المحلى ، وأن عددا جد قليل من الأفراد المحليين ، المدرسين فى المدرسة مثلا ، لديهم الوقت ، أو المهارة لانشاء اختبارات يمكن استخدامها فى نطاق واسع ، هذه الحقيقة تحد كثيرا من الامكان والقيمة التطبيقية للاتجاه الى بناء اختبارات تحصيلية مرجعة الى المحك ٠

وعلى اية حال ، فى الولايات المتحدة ، نجد فى زيادة « بنوك » أو « مخازن » مرامى الأداء performance objectives ، وعنــاصر الاختبار وبنوده ٠ وزيادة رغبة مديرى المدارس فى أن يكونوا قوائم مرامى الأداء لمناطقهم ٠ ونمو شركات الاختبار فى هذا المجال الخاص بالاختبارات المرجعة الى المحك ، كل هذا زاد من احتمال ان الاختبارات المرجعة الى المحك سوف تكون ميسرة للاستعمال فى قياس التحصيل المدرسى ٠

ولهذه لاختبارات المحكية ، سيما يبدو ، مزية السماح لكل منطقة تعليمية بأن توجه برنامجها التقويمى نحو أهدافها الخاصة ، وأن تعرف مدى احراز هدفها بشكل مطلق وليس بشكل نسبى ٠ فليس كافيا أن تقول ان محمدا تعلم اكثر من حامد ، ذلك لأنه مازال علينا أن نعرف عدد المرامى التى حققها كل منهما حتى تسمح بالتقدم نحو أهداف أخرى اكثر تركيبا ٠ وسوف يقــدم ناشرو الاختبارات مساعدة كبيرة للمدارس اذا أمدوها بعناصر الاختبار التى

تتفق مع المرامى بحيث تكون كل مدرسة قادرة على أن تشـــكل اختبارها التحصيلى وتصوغه ٠ موجها نحو حاجات تلاميذها ٠

٥ ـ تفسير درجات الاختبار المرجع الى المحك

درجة الاختبار ليست سوى رقم ٠ ماذا تحمل هذه الدرجة من معنى ودلالة ؟ ماذا تكشف عنه ؟ هل أبدى التلميذ كفاءة فى المرامى المقاسة ؟ هل لديه قدرة ليتحرك الى مستوى أداء تال ؟ هل لديه القدرة على أداء مهارات على مستوى مطلوب لبداية العمل فى مهنة ما ؟

درجة الاختبار فى ذاتها لا تدلنا الا على القليل جدا ، وحتى تكون مفيدة يجب أن يعطى لها تفسير ٠

واذا كان الاختبار قائما على عدد من المرامى ، وكانت هذه المرامى ذاتها لها درجة من الصدق ، فان الدرجة على الاختبار سوف تخبرنا بشىء عن قدرة التلميذ على الأداء فى مجال الاختبار ٠ ويجب أن يمدنا الاختبار المرجع الى المحك بمعلومات عن درجة كفاءة التلميذ فى المرامى التى يقيسها الاختبار ٠ فالدرجة المرتفعة على الاختبار يجب أن تدل على احراز المرامى وتحصيلها ، بينما الدرجة المنخفضة تدل على نقص التحصيل ٠ وقد تظن إنه ربما كانت عناصر الاختبار سهلة جدا او ضعيفة جدا ، ومن ثم فانها تكون غير منصفة فى قياس التحصيل ٠ ولكن اذا كان اختبارك قد استوفى محكات الملاءمة ، والصدق ، والثبات ، فانه من الأرجح ان تكون عناصره مناسبة لقياس التحصيل ، ومن ثم يمكنك ان تحكم بأن النجاح فى الاختبار يقيس احراز المرامى وتحصيلها ٠

وحتى يكون تفسير درجة الاختبار التحصيلى دقيقا وسليما ، ربما يكون من الأفضل ان تقيد نفسك بالدرجة الكلية على الاختبار ٠ ولما كان الاختبار المرجع الى المحك اكثر شبها بمجموعة من الاختبارات المصغرة mini-tests يقيس كل منها مرمى واحدا ٠ وكل منها يحتوى عنصرين ، وأحيانا اكثر من عنصرين ، لما كان هذا شان الاختبار المرجع الى المحك ، فان التحصيل يمكن ان يوصف عن طريق ذكر المرامى الخاصة التى اظهر فيها كل تلميذ كفاءة

بحكم نجاحه فى الاختبار المصغر المناسب او المتصل بهذه المرامى · وعلى هذا ، فانه لابد من ان نحتاج الى أن نقرر ما هى الدرجة المرتفعة التى يمكننا اعتبارها مرتفعة · يمكن أن نسجل لكل تلميذ اسم المرمى الذى اظهر فيه كفاءة ، ولا بأس - اذا كان هذا مرغوبا فيه - باعطاء التلميذ ما يدل على مستوى الكفاءة الذى احرزه فعلا · فتفسيرية مثل هذا الاختبار قائمة اساسا لا على الدرجة الكلية بل على دلائل درجة الكفاءة التى امكن احرازها لكل مرمى من المرامى التى يقيسها الاختبار · ولما كانت المرامى تمثل اهتمامنا الحقيقى ، وكانت الاختبارات المرجعة الى المحك تقيس تحصيل هذه المرامى فاننا نستطيع أن نركز نتائج اختبارنا وتفسيراتنا على ذلك التعلم الذى نريد أن ندفعه ونرفعه ·

الاختبارات التحصيلية المرجعة الى المحك

الترجيع الى المعيار norm-referencing واحد من أهم مكونات الاختبارات المقننة · والاختبارات التحصيلية المرجعة الى المعيار تفسر درجاتها بمقارنتها بدرجات حصل عليها بالفعل تلاميذ اخرون · أولئك الذين يكونون الجماعة المعيارية أو جماعة المعيار norming group وتقرر ما اذا كانت درجة ما منخفضة أو مرتفعة بمقارنتها بدرجات حصل عليها تلاميذ من مستوى الصف ذاته فى الجماعة المعيارية ·

ويمكن القول ان كل اختبار ينشئه المدرس هو بالامكان اختبار مرجع الى المحك ، لأنه مصمم لقياس درجة الكفاءة المنخفضة على عدد معين من المرامى · ومن العجيب أن كثيرا من المدرسين لا يبنون اختباراتهم حول مرام معينة ومحددة ، بل ان كثيرين منهم يصححون اختباراتهم كما لو كانت مرجعة الى المعيار بأن يقيموا تصحيحهم على اساس الأداء النسبى للمفحوص ، فى حين انه يجب تصحيح الاختبار فى ضوء أداء على مرام أو عناصر ، اى على اساس الترجيع الى المحك ·

الاهتمام الرئيسى فى الاختبارات المرجعة الى المعيار هو : الى اى مدى يمكن أن يقارن فرد أو مجموعة افراد مع آخرين ؟ والاجابات عن هذا السؤال

مفيدة جدا فى ضمان تحقيق مستوى ادنى من الأداء النسبى فى الفصل ، أو المدرسة ، أو المنطقة التعليمية ٠

اما الاختبارات المرجعة الى المحك فتسأل : بأى شكل يسـلك فرد أو مجموعة من الأفراد ؟ ما الذى يعرفه الفرد ذاته ؟ ما الذى تعرفه المجموعة ؟

ومن ثم فان الترجيع الى المعيار يعطي نتائج تعيينيـة ، summative أى انها تخبرنا اين يقف الفرد أو مجموعة الأفراد ، بينما الترجيع الى المحك يعطينا نتائج تشكيلية formative ، أى أنها تخبرنا فى أى المجالات يمكن أن يوجه التدريس بحيث يتيسر تحصيل الكفاءة ٠ الاختبارات المرجعة الى المحك تساعد المدرسة على تعرف monitor تقـدم التلميـذ ٠ وتشخيص مواضع القوة والضعف ، ووصف التدريس وتوجيهه ، وخاصة العـلاجى منه ٠

وبينما يستطيع المدرسون انشاء اختباراتهم التحصيلية المرجعة الى المحك ، فان الاختبارات المنشورة تعطيهم درجة أرفع من ضـبط نوع اداة الاختبار ، وامكانات افضل فى ملامح تقرير نتائج الاختبـار ، وفى امكان المدرسين أن يفسروا هذه النتائج فى ضوء مراميهم التدريسية الخاصة ٠

ومن الممكن بناء اختبار تحصيلى مقنن ، معيارى ، يكون فى الوقت ذاته اختبارا مرجعا الى المحك ٠ كيف ؟

مقارنة بين الاختبارات المرجعة الى المعيار ، والاختبارات المرجعة الى المحك

تعرف أن الاختبارات التحصيلية المقننة يتوافر لها ثلاث خصائص هى :

١ ــ أن عناصرها وبنودها جربت ، وحللت ، وروجعت

و ٢ ــ انتشارها واتساع استعمالها واعادة استعمالها المقنن ٠

و ٣ ــ وجود معايير لها واستعمال هذه المعايير فى تفسيرها ٠

وانه من الممكن بناء اختبار تحصيلى مقنن يكون فى الوقت ذاته مرجعا

الى المحك ، وذلك بتغيير واحدة من الخصائص الثلاث السابقة ، تلك هى الخاصية رقم (٣) ، بحيث تكون هكذا ·

« وجود مرام للتفسير واستعمالها ، وتتضمن هـــذه المرامى محكات لتقويم الكفاءة ، فعندئذ نكون قد وضعنا الأساس · على الأقـــل من حيث التعريف ، لاختبار تحصيل « مقنن » ، يكون اختبارا تحصيليا مرجعا الى المحك ·

هذا ، وان الفرق الوظيفى بين اختبار مرجع الى المحك (رك CR) وآخر مقنن أو مرجع الى المعيار (رع NR) ذو ثلاث شعب :

الأولى : أن الاختبار المرجع الى المحك موجه توجيها خاصا الى مرام بينما الاختبار المرجع الى المعيار تكون درجاته اكـــثر كلية more global وشمولا ، ومن ثم فانك تحصل على درجات اكثر فى الاختبـــار المرجع الى المحك ، (لأنك تحصل على درجة واحدة لكل مرمى) ، ولكن يوجد فيه عدد أقل من العناصر لكل درجة ·

والثانية : درجات الاختبار المرجع الى المعيار يمكن أن تفسر عن طريق المعايير التى يعطيها ناشروه ، بينما درجات الاختبار المرجع الى المحك تتطلب منك أن تضع أنت نقطة أو درجة القطع التى تدل على الكفاءة المناسبة ·

والثالثة : عناصر الاختبار المرجع الى المعيار تكتب بحيث تنتج أقصى قدر من التنوع فى الأداء بين التلاميذ ، بينما الاختبـــار المرجع الى المحك تكتب عناصره بحيث تمثل نطاق المرامى كما يراه كاتب العنصر ·

بعد أن نتعمن فى هذه الفروق بين الاختبار المعيارى NR (رع) والاختبار المحكى (رك) ، وبعد معرفة أن عناصر الاختبار المعيارى قد كتبت مستعملة تخطيط محتوى content outline أو تصنيف مهارات skills classification ، فمن الممكن أن نفكر فى استخدام اختبار معيارى كما لو كان اختبار محك على شرط أن يذكر ناشره نتائجه بمجـــال المحتوى أو مجال المهارة · وقد بدأ كبار ناشرى الاختبارات منذ الآن يفعلون

هذا بالنسبة للاختبارات المعيارية أو المقننة مثل اختبار أيوا Iowa Test
Metropolitan Achievement Tests of Basic Skills .

وان تفسيرات نتائج الاختبارات التحصيلية المرجعة الى المحك يغلب أن يكون مساعدا للمدرسين بصورة اكبر من تفسيرات الاختبارات المرجعة الى المعيار ، الا أننا يجب أن ننتبه الى نقطتين :

الأولى : من الضرورى أن تكون المهارات والمحتويات المقاسة على وجه الخصوص هى تلك التى تحتاج معلومات عنها ، وأنه يوجد عدد كاف من العناصر لكل مهارة ، وذلك لتوفير الثبات .

والأخرى : لما كانت العناصر مكتوبة لأقصى تمييز بين الدرجات المرتفعة والدرجات المنخفضة ، فانك قد تحتاج الى فحص درجـــــات تلاميذك على الاختبار لتحديد أى نقطة قطع ، أو درجة قطع ، تستخدمها للكفاءة المناسبة بالنسبة الى كل مهارة ، بدلا من استخدام نقطة قطع ثابتة مقررة . وبمعنى ما فان هذا مازال صورة من صور الترجيع الى المعيار ، ذلك لأن اختبار نقط القطع سوف يقوم على الأداء النسبى relative performance للتلاميذ ، مثلا ، عدد العناصر المجابة صوابا من ٧٠٪ من التلاميذ .

مراجع الباب الخامس

١ ـ احمد زكى صالح : الأسس النفسية للتعليم الثانوى ، دار النهضة العربية ، ١٩٧٢ .

٢ ـ رايتسون ، ج . واين واخرون (ترجمة محمد محمد عاشـور واخرين) : التقويم فى التربية الحديثة . مكتبة الأنجلو المصرية ١٩٦٥ .

٣ ـ رمزية الغريب : التقويم والقيـــاس النفسى والتربوى . الأنجلو المصرية ، ١٩٧٠ .

٤ ـ سامى ابراهيم على : تقنين امتحانات البيانو . رسالة ماجستير ، المعهد العالى للتربية الموسيقية ، ١٩٧١ .

٥ ـ فان دالين ، ديوبولد (ترجمة محمد نبيل نوفل واخرين) : مناهج البحث فى التربية وعلم النفس : مكتبة الأنجلو المصرية ، ١٩٦٩ .

٦ ـ فؤاد أبو حطب : دور التربية فى تنمية التفكير الابتكارى . مجلة الفكر المعاصر ، يونيو ١٩٧٠ .

٧ ـ ميكر ، روبرت ف . (ترجمة جابر عبد الحميد جابر وسعد عبد العزيز نادر) : الأهداف التربوية ، بغداد ، مطبعة العانى ، ١٩٦٧ .

٨ ـ نعيم عطية : التقييم التربوى الهادف ، بيروت ، منشورات دار الكتاب اللبنانى ، ١٩٧٠ .

9. Bloom. B. Taxonomy of Educational Objectives. London : Longmans, 1956.

10. De Cecco, J.P. The psychology of Learning and Instruction. Englewood Cliffs, N.J., 1968.

11. Mager, R.F. Preparing Objectives for Programmed Instruction. Palo Alto, Calif. : Varian Associates, 1961.

12. Ebel, B.L. Maximizing test validity in fixed time limits. Educ. psychol. Measurements, 1953, 8, 347-357.

13. Ebel, R.L. Measuring Educational Achievement, Prentice-Hall, 1965.

14. Gagne, R.M. Conditions of Learning, Holt., Rinehart, Winston, 1965.

15. Gronlund, N.E. Measurement and Evaluation in Teaching, Macmillan, 1965.

16. Linquist, E.F. (ed.) Educational Measurement. Amer. Council on Educ., Washington, D.C. 1951.

17. Lindvall, C.M. Testing and evaluation : An Introduction. Harcourt, Brace, 1961.

18. Mager, R.F., Mc Cann, J. Learner-Controlled Instruction Palo Alto, Clif Varian Associates, 1961.

19. Thorndike, R.L. (ed.) Educational Measurement (2nd ed.) Amer. Council on Educ., D.C., 1971.

البـاب السادس

مقاييس الشخصية

الفصل الثالث عشر

وسائل التقرير الذاتى

اشرنا فيما سبق الى تقسيم المقاييس النفسية بصفة عامة الى قسمين رئيسين هما : مقاييس الأداء المميز ، ومقاييس اقصى الأداء ، وعرفنا ان فى النوع الأول من هذه المقاييس يكون الهدف من القياس هو معرفة ما يقوم به الشخص عادة ، بينما النوع الآخر من المقاييس يكون الهدف هو معرفة اقصى ما يمكن ان يقوم به الشخص فى ظروف استثارة ودافعية غير عادية ٠

ويمكن أن نضع تحت القسم الأول ، أى مقاييس الأداء المميز ، اختبارات الشخصية ، واختبارات الميول والاتجاهات بينما تدخل اختبارات القــدرة بانواعها المختلفة تحت القسم الثانى وهو اختبارات اقصى الأداء ٠ وفى هذا الفصل نتكلم عن قياس الشخصية باعتباره قياسا للسلوك أو الأداء المميز ٠ وسوف نتبع ٠ بصفة عامة ، التقسيم الذى حدده كرونباك (٢٧) لأنواع مقاييس الشخصية ، فندرجها تحت هذه الأقسام الثلاثة :

(أ) قياس الشخصية عن طريق أحكام الآخرين والملاحظات المنظمة Judgments and systematic observations ، وهى ما سيتناوله الفصل الرابع عشر ٠

(ب) قياس الشخصية عن طريق اختبارات الأداء performance tests وهى ما سيتناوله الفصل الخامس عشر ٠

(ﺟ) قياس الشخصية عن طريق التقــرير الذاتى self-report وهى ما سيتناوله الفصل الحالى ٠

وقياس الشخصية عن طريق التقرير الذاتى Self-report على الرغم من أن تعبير السلوك المميز الذى تكررت الاشارة اليه يحمل طابعا «سلوكيا» فان المذهب السلوكى لم يكن هو الذى اسهم فى نشــأة وتطوير مقــاييس الشخصية وانما يرجع الفضــل فى هذا الى الاتجــاه الظاهراتى

(الفينومونولوجى phénomenology) ، فهذا الاتجـــاه يتميز بانه يركز على العالم، كما يدركه الشخص ، أو على العالم الذاتى للشخص ، بما فى ذلك ادراك الفرد لذاته ، وعلى ان السلوك انما يتحدد ، ولا يمكن ان يفهم الا فى ضوء هذا العالم الذاتى الداخلى ، كما يذهب هذا الاتجاه الى ان غالب مشكلات التوافق التى يتعرض لها الفرد ترجع الى ادراك الشخص للاحداث والأشخاص وليس لواقع هذه الاحداث او الأشخاص فى ذاتهم • وبالتالى فان العلاج النفسى ، وفق هذا الإتجاه يجب ان يركز على تصحيح ادراك الشخص •

نشير الى المدرسة الفينومونولوجية فى بداية كـــلامنا عن قيــاس الشخصية لأن أول مقاييس او استفتاءات للشخصية كانت قد وضعت بهدف قياس هذا العالم الداخلى الذاتى من الادراك والشعور • ففى العقد التاسع من القرن التاسع عشر وضع جالتون Francis Galton هذه المقاييس عندما احتاج الى طريقة يمكن أن يطبقها على عدد كبير من مفحوصيه اثناء دراساته للصور العقلية ، ثم استعمل هذا النوع من الاستفتاءات بعد ذلك ستانلى هول Stanely Hall فى دراساته الموسعة عن نمو المراهقين ولكن الاستفتاء كانت له وظيفة خاصة بالنسبة الى كل من جالتون وستانلى هول • فبالنسبة الى جالتون كان التقرير الذاتى يستعمل على انه الوسيلة الوحيدة الممكنة الحصول على معلومات عن أمور واحداث فى عقل المفحوص ، فى داخـــل راسه ، بينما كان التقرير الذاتى بالنسبة الى ستانلى هول يستعمل بهـدف تجنب الجهد الشاق والوقت الطويل الذى يمكن أن يستهلك فى الملاحظـــة المباشرة لسلوك المراهقين (٢٧ : ٤٦٤) •

اما أول استفتاء وضع بقصد قياس الشخص فقد كان قائمة وود ورث للبيانات الشخصية • Woodworth Personal Data Sheet وظهرت خـــلال الحــرب العالمية الأولى عنـدما ظهـرت الحاجة الى تمييز اولئك الجنود الذين عندهم قابلية للانهيار اثناء القتال ، ولم يكن يتوافر عدد كاف من المتخصصين لتحقيق هذا الهدف • لهذا وضع وود ورث قائمة من الاعراض التى يهتم الفاحص بالتعرف عليها لاستبعاد من تظهر عنـدهم ، وعندما أعطيت القائمة ميزت بالفعل أولئك الذين قرروا وجود عدد كبير من

الأعراض عندهم ثم خضعوا لاختبار اكثر دقة ، وكانت لاختبـــار البيانات الشخصية هذه قيمة كبيرة فى ذلك المجال ·

ويعتبر اختبار وود روث الجد الأكبر لعدد كبير من استفتاءات التوافق ، وهى التى تتكون اساسا من قوائم من المشكلات او الأعراض او الشكاوى التى على المفحوص أن يختارها · الا ان مثل هذه القوائم لا تدعى انها تعطى وصفا دقيقا عن الشخصية وهى غالبا ما تعطى درجة واحدة تمثل مستوى التوافق · وهى تتكون من عناصر يمكن أن تميز بين اشخاص معروف انهـم غير متوافقين وبين آخرين معروف انهم عاديون ، والاســـتعمال الرئيسى لمثل هذه المقاييس هو التعرف على أولئك الأشـــخاص الذين يحتاجون الى رعاية وارشاد نفسى ، وعلى هذا يمكن أن ننظر اليها على انها ادوات فرز screening تساعد فى تحديد الأشخاص الذين يختارون اعراضا وصفات وانتقادات تتصل بهم شخصيا ، وقد عدله بيرت وعربه واستعمله احمد زكى صالح (١ : ٨٥١ ـ ٨٥٢) ·

ولا يمكن اعتبار استفتاءات التوافق مقاييس دقيقة لسمات محددة ، فالمعلومات التى تمدنا بها معلومات ليس لها صفة العمق · لهذا فانه فى الفترة بين ١٩٢٠ ـ ١٩٤٥ · تغيرت النظرة الى مثل هذه الاستفتاءات وذلك بتاثير المدرسة السلوكية واستنكارها للتأمل الباطنى الذاتى الذى هو أساس اختبارات التوافق · فيجب ان يكون المقياس بديلا لملاحظة السلوك ، وعلى هذا يجب ان تؤكد اسئلته بشكل اكبر على ما يفعله الشخص لا على ما يشعر به او يفكر فيه · وعلى هذا اتسعت مجالات الاختبارات بحيث تصف أكبر عدد ممكن من جوانب الشخصية · وتلخص الاستجابات باعطاء درجـــات لعدد من السمات traits او لأنماط الاستجابات ، وذلك لأنه كان ينظر الى الشخصية فى ذلك الوقت على انها مجموعة او حزمة من العادات ، وكان يحكم على الشخص حسب قوة سمات مثـــل : الثقة بالنفس ، الاصرار ، سهولة تكوين صداقات ·· الخ والسمة القوية هى السمة التى تصف استجابة يقوم بها الشخص عادة وبكثرة ·

وقد كانت قائمة السمات او اقسام السلوك التى تعطى لها درجات فى الاختبارات تختار بشكل قطعى ودون خطة معينـــة فى الاختبارات الأولى ،

فمثلا سمة مثل الثقة بالنفس كان مصدرها الخبرة اليومية ، بينما تشتق سمة أخرى مثل الانطواء من نظرية من نظريات الشخصية · وقد ظهرت عشرات من الاختبارات بهذا الشكل ، تأخذ اللاحقة منها من السابقة عليها بضعة عبارات أو أسئلة وتضيف اليها وتصممها بحيث توضع فى تجمعات سمات مختلفة · الا أن أكثرها شيوعا هو اختبار بيرنرويتر للشخصية ·
Bernreuter Personality Inventory.

وقد ترجمه الى العربية محمد عثمان نجاتى (*) وتعطى فيه الدرجات للدلالة على : النزعة العصابية ، الاكتفاء الذاتى · الانطواء ، والسيطرة · ويتكون الاختبار من ٢٥ سؤالا يجيب عنها المفحوص بنعم أو لا · وقد وضعه بيرنرويتر على أساس أن سلوك الفرد فى موقف واحد قد يكشف عن سمات شخصية مختلفة ، وعلى هذا يمكن أن ينظر الى كل سؤال من أسئلة الاختبار على أنه يقيس أكثر من سمة من السمات ، وبالتالى يمكن أن تستخدم الاختبار الواحد للكشف عن أكثر من سمة تبعا لطريقة معالجة الاستجابات التى تحصل عليها (٤) ·

وقد كانت دراسة فلاناجان (٢٩) Flanagan لاختبار بيرنرويتر بداية لاستعمال درجات السمات كما تحددها القواعد الاحصائية · لقد اعتقد فلاناجان أن السمات لا يمكن أن نطلق عليها أسماء خاصة بها الا اذا كانت الارتباطات فيما بينها منخفضة · وعلى هذا قام بحساب معاملات الارتباط لدرجات اختبار بيرنرويتر من تطبيقه على ٣٠٥ مراهق ، ووجد أن هذه السمات ليست مستقلة · فقد وجد أن الانطواء لا يختلف الا قليلا عن النزعة العصبية ، وعندما طبق التحليل العاملى وجد عاملين هما : الثقة بالنفس والميل الاجتماعى يمكن أن يفسر البيانات التى نحصل عليها من قياس السمات الأربعة الأصلية ، وأنشأ مفاتيح خاصة لقياس العاملين أو السمتين اللتين وصل اليهما أى الثقة بالنفس والميل الاجتماعى · وقد ارتبطت درجات هاتين السمتين ارتباطا ضعيفا جدا ، وعلى هذا أصبح من الممكن أن نقول أنهما يمثلان جانبين مستقلين من التقرير الذاتى ·

(*) لا تتوافر لدينا بيانات سيكومترية عن هذا الاستخبار فى البيئة المصرية ·

بعد هذا حدث تطور آخر له مغزاه ، ذلك هو اخضاع نظريات الشخصية تماما لبحث احصائى عن ابعاد dimensions يمكن أن تلخص الشخصية . فقد أدت الارتباطات الداخلية بين عناصر الاختبارات الى أن اقترح جيلفورد مثلا ، أن الانطواء يمكن أن نقسمه الى انطواء اجتماعى (ج) ، وانطواء فكرى (ف) ، واكتئاب (ك) ، والتقلب المزاجى (ت) ، وضبط للنفس (ض) . وعلى هذا انشأ الاختبار الذى أسماه : اختبار العوامل ج - ف - ك - ت - ض Inventory of Factors S - T - D - C - R.

ثم أضاف بعد هذا ثمانية جوانب أخرى للشخصية . ولم تكن مقاييس جيلفورد مستقلة ، ولهذا أعاد بعض الباحثين تنظيم هذه المقاييس فى أنماط تجعلها أكثر كفاءة . من هؤلاء مثلا ترستون الذى استطاع أن يفسر كثيرا من المعلومات فى مقاييس جيلفورد الثلاثة عشرة بسبعة عوامل فقط وأعاد تسميتها بحيث أصبحت هكذا التأمل ، الميل الاجتماعى ، الاتزان الانفعالى ، الحيوية أو الذكورة ، السيطرة ، النشاط ، الاندفاع ، وقد اقتبس منه أحمد زكى صالح اختبار الصفات الانفعالية .

وهكذا نجد أن واحدا من علماء النفس يصنف عناصر الاختبار التى تقيس السمات بشكل دقيق تفصيلى ، بينما يأتى التالى ليضعها فى مجموعات صغيرة . ثم يأتى آخر فيعيد تقسيم المجموعات بطريقة جديدة ، ويعطى كلا منها أسماء جديدة للعوامل التى يصل اليها . وبهذا لا تنتهى هذه اللعبة ، كما يعبر كرونباك (٢٧ : ٤٦٧) عن رأيه أنه ما لم ترتبط قوائم السمات بنظرية محددة وبمحكات خارجية ، فان عملية الاختبار لا يمكن إلا أن تكون عملية تفضيلية شخصية . وليس هناك اتفاق فى الوقت الحاضر بين أصحاب التحليل العاملى ، بالنسبة الى عدد العوامل التى تم تحديدها بشكل دقيق ، أو افضل تنظيم لها ، أو انسب الأسماء لها .

أما الاتجاه الأحدث فى قياس الشخصية بصفة عامة فهو تحديد المفاهيم والتكوينات النظرية Constructs على أساس نظرية للشخصية تم اعداد العناصر فى الاختبارات بصورة تؤدى الى الحصول على معلومات عن هذه المفاهيم والتكوينات . وقد كان لنظرية يونج Jung عن الشخصية تأثير فى الدراسات الأولى عن الانطواء . ولكن هذه النظرية لم تتعد اقتراح بعض

عناصر اختبارات وضعت لمجرد المحاولة ٠ أما الآن فهناك دراسات واسعة عن اختبار مايرز - برجز Myers-Briggs Inventory والتى اشتقت عناصره ومفاتيح تصحيحه صراحة من نظرية يونج (٥١) ٠ وتوجد مقاييس أخرى تمثل هذا الاتجاه منها مقياس التفضيل الشخصي لادوادر Edwards Personal Preference Schedule.

والذى يشتق من نظرية مورى Murray عن الحاجات الخمس عشرة ٠ ومن أمثلة هذه الحاجات : الانجاز ، والخضوع ، والنظام ، والعرض ، والاستقلال الذاتى ، السيطرة ، التأمل النفسى (المقصود بهذه الحاجة تحليل دوافع ومشاعر الفرد الذاتية وعند الآخرين) ، ويتكون المقياس من ٢١٠ زوجا من العبارات ، وعلى المفحوص أن يختار عبارة من كل زوج على انها أكثر تمثيلا لخصائص شخصيته ، ومن أمثلة هذه الأزواج ٠

(أ) أحب أن أتحدث عن نفسى مع الآخرين ٠
(ب) أحب أن أعمل لهدف معين حددته بنفسى ٠

(أ) أشعر بالاكتئاب عندما افشل فى أمر ما ٠
(ب) أشعر بالاضطراب عندما اتحدث الى جماعة ٠

ويهدف هذا المقياس الى تقدير القوة النسبية للخمسة عشر متغيرا المرتبطة بالحاجات ٠

ويمكن تقويم الخمسة عشر درجة المرتبطة بالحاجات على أساس معايير مئوية وتائية لطلاب وطالبات الجامعة ، وهذه المعايير تقوم على عدد ٧٤٩ طالبا و ٧٦٠ طالبة موزعين على جميع أنحاء الولايات المتحدة ، وقد أضيفت معايير مئوية للكبار ينتمون الى بيئات حضرية وريفية ٠

أما عن ثبات المقياس فقد ذكر دليل الاختبار معاملات ثبات لمقاييسه الفرعية الخمس عشرة معتمدة على اعادة الاختبار بعد اسبوع واحد تتراوح بين ٧٤ر٠ ، ٨٨ر٠ ، بينما معاملات الثبات المعتمدة على التجزئة النصفية تتراوح بين ٦٠ر٠ ، ٨٧ر٠ ٠

أما بالنسبة الى صدق مقياس ادواردز ، فعلى الرغم من قلة البيانات الواردة عنها فى دليل الاختبار ذاته ، فان ابحاثا كثيرة مستقلة حول صدقه قد نشرت · الا انها أعطت نتائج متعارضة ، وتذهب انستازى (١٨ : ٤٥٤) الى أن مقياس التفضيل الشخصى له مزايا كثيرة الا أنه محتاج الى أمرين :

اولا ـ مراجعة للتخلص من بعض جوانب الضعف الفنية فيه ، وبصفة خاصة فيما يتصل بشكل العناصر وتفسير الدرجة ، وبالثانى ، دراسات صدق تجرى بدقة مستفيدة من أساليب تحليل أنماط الدرجات مناسبة لدرجات هذا المقياس التى لا تتحدد فيها قوة الحاجة بصورة مطلقة ولكن بالنسبة الى قوة الحاجات الأخرى عند الفرد ، وعلى هذا فان اطار المرجع فى هذا النوع من الدرجات هو الفرد ذاته وليس عينة التقنين : وقد يكون شخصان متفقان فى درجاتهم على المقياس الا أنهما يختلفان اختلافا جوهريا فى القوة المطلقة لحاجاتهم ·

وقد عرب هذا المقياس جابر عبد الحميد (١٧ : ٤٧٨ ـ ٤٨٠) وحسب ثبات الاختبار بطريقة التجزئةالنصفية، وذلكبتطبيقه علىطلاب الجامعة. وكان يتراوح بين ٣٤ر٠ ، ٧٧ر٠ وقد استعمل هذا المقياس فى بعض الدراسات المقارنة بين الثقافات وغيرها ·

ومن المقاييس الأخرى التى اعتمدت على نظرية مقياس تيلور للقلق الظاهر Tayler Manifest Anxiety Scale وقد ارتبط فى تصميمه بالبحث فى نظرية هل ـ سبنس Hull-Spence السلوكية ، وسوف نعود اليه بعد دراستنا لاختبار مينسوتا المتعدد الأوجه الذى توليه اهتماما خاصا·

اختبار مينسوتا المتعدد الأوجه للشخصية :

Minnesota Multiphasic Personality Inventory (MMPI)

يعتبر هذا الاختبار من أهم الأدوات التى وضعت لقياس الشخصية عن طريق التقرير الذاتى ، لهذا سوف نتناوله بشىء من التفصيل · نجد لهذا

الاختبار مكانا بين اختبارات الشخصية تشبه مكان اختبار سترونج Strong بين اختبارات الميول ، ذلك لأنه يشارك اختبار سترونج فى الأساس التجريبى الذى صمم وفقا له ، وكذلك يشبهه من حيث تعرضه لبحث عميق وطويل من جانب واضعيه . وقد ظهر هذا الاختبار فى وقت كانت الحاجة اليه شديدة ، وكان الاعتماد عليه كبيرا اثناء وبعد الحرب العالمية الثانية التى شـــهدت توسعا كبيرا فى علم النفس المرضى والعلاج النفسى . اى أن نشأته كانت أصلا نشأة اكلينيكية وهذا مما ساعد على انتشاره ، وكذلك ساعد على انتشاره ايضا بعض الأمور الفنية المبتكرة فيه ومنذ نشره فى عام ١٩٤٠ ونشر دليله فى عام ١٩٤٣ فانه استثار فيضا من البحوث التى استخدمته .

وكان صميم الاختبار فى الأصل على يد عالم نفس هو ستارك هاثاوى Starke Hathaway وطبيب نفسى هو ج.س. ماكنلى J.C. Mckinley (٣٥) وكان الهدف من هذا التصميم هو المساعدة فى تشـــخيص المرضى النفسيين وتحليل السمات التى يميز الشذوذ النفسى الذى يؤدى الى العجز.

يتكون الاختبار من ٥٥٠ عبارة يسال المفحوص أن يصنفها وفق أحـد اقسام ثلاثة ، اى يجيب عنها اما بنعم أو لا ، ولا استطيع أن اجزم أو لا أدرى وفى الصورة الفردية للاختبار كانت العبارات تطبع على بطاقات منفصلة وعلى المفحوص ن يصنفها فى ثلاث مجموعات ، وبعد ذلك طبعت العبارات فى كتيب وتسجل الاستجابات فى ورقة الاجابة . وقد صممت الصـــورتان الفردية والجماعية للكبار ، اى من ١٦ سنة وما بعدها . الا انهما استخدما بنجاح مع مراهقين اصغر من هذا السن . وبصفة عامة يفضل استخدام الصورة الفردية وخاصة عند الأشخاص المضطربين أو من ذوى المستوى التعليمى أو العقلى المنخفض .

وتتنوع عناصر الاختبار المتعدد الأوجه للشخصية وتتسع من حيث محتواها فتشمل مجالات مثل . الصحة ، والأعراض الجسمية - النفسية ، والاضطرابات العصبية ، والاتجاهات نحو الجنس ، والسياسة ، والدين ، والمسائل التربوية والمهنية والعائلية ، كما تشمل كثيرا من المظاهر والأعراض العصابية والذهانية المعروفة مثل المخاوف المرضية ، والاضطهاد ، ونزعات تعذيب الذات وتعذيب الآخرين وغيرها .

وعندما نشر الاختبار فى صورته الأولى كان يعطى درجات على تسعة « مقاييس كلينيكية » ، ويتكون كل مقياس منها من عناصر ميزت بين جماعة مرضية محدودة وبين جماعة ضابطة عادية ، وعلى هذا فان هذه المقاييس قد انشئت بطريقة تجريبية ، وذلك عن طريق تقسيم العناصر وفق محك معين وهو التشخيص النفسى المعروف ووفق هذه الطريقة اعدت المقاييس التالية :

1. HS : Hypochondriasis	يسمى	١ ـ توهم المرض ـ هـ س
2. D : Depression	ثلاثى	٢ ـ الاكتئاب ـ د
3. Hy : Hysteria	العصاب	٣ ـ الهستيريا ـ هـ ى
4. Pd : Psychopathic deviation		٤ ـ الانحراف السيكوباتى ـ ب د
5. Mf : Masculanity-femininity		٥ ـ الذكورة ـ الأنوثة ـ م ف
6. Pa : Paranoia		٦ ـ البارانويا ـ ب ا
7. Pt : Psychasthenia		٧ ـ السيكائينيا ـ ب ت
8. Sc : Schizophrenia		٨ ـ الفصام ـ س ك
9. Ma : Hypomania		٩ ـ الهوس الخفيف ـ م ا

وفيما يلى تعريف مختصر بما يدل عليه كل من هذه المقاييس (١٥ : ١٥٢ ـ ١٦٥) ٠

١ ـ توهم المرض : وهو مقياس للاهتمام الزائد بالوظائف الجسمية ، والقلق الذى لا مبرر على الصحة ٠

٢ ـ الاكتئاب : ويرتبط هذا المقياس بالشعور بالانقباض والتشاؤم والاحساس بعدم الأهمية الذاتية ٠

٣ ـ الهستيريا : ويقيس هذا المقياس درجة تشابه المفحوص بالمرضى الذين تظهر عندهم أعراض تتخذ صورة شكاوى عامة أو محددة مثل العمى أو الشلل أو القىء أو الاضطرابات المعوية ٠ أو الأعراض القلبية ٠

٤ ـ الانحراف السيكوباتى : وهو مقياس لدرجة تشابه المفحوص بجماعة السيكوباتيين الذين تتمثل الصعوبة الرئيسية عندهم ، فى نقص

الاستجابة الانفعالية العميقة وقلة الاهتمام بمشاعر الآخرين ، وعدم المبالاة بالمعايير الاجتماعية ٠

٥ ـ الذكورة ـ الأنوثة : وهو مقياس للذكورة والأنوثة فى انمـــاط الاهتمامات ٠ وقد استخرجت عبارات هذا المقياس نتيجة مقارنة استجابات الذكورة من ذوى الاهتمامات الذكرية باستجابات الذكور من ذوى الاهتمامات الأنثوية ، اما الفروق بين استجابات الذكور والاناث فقد درسها واضـــعا الاختبار الا أن هذه الفروق لعبت دورا صغيرا نسبيا فى هـــذا المقياس ٠ وتدل الدرجات المرتفعة على هذا المقياس على غلبة انماط الاهتمامات الخاصة بالجنس الآخر ٠

٦ ـ البارنوايا : وهذا مقياس للتشكك الزائد ، والحساسية المفرطة ، وهواجس الاضطهاد ٠

٧ ـ السيكاثينيا : وهو مقياس للمخاوف المرضية والسلوك القهرى ، الذى قد يكون صريح مثل التكرار الزائد لغسل اليدين ، أو ضمنيا كما يتمثل فى عدم القدرة على التحرر أو الهروب من فكرة متسلطة ٠

٨ ـ الفصام : ويكشف هذا المقياس عن التفكير أو الســـلوك الخلطى الشاذ والذى لا يتفق مع الاتصال بالواقع وفهمه ٠

٩ ـ الهوس الخفيف : وهو مقياس للنشاط الزائدة فى الفكر والعمل ، وعدم التركيز الطويل ، والانشغال بعدد كبير من الأمـــور والمشروعات فى وقت واحـــد ٠

وتعطى بيانات جماعة التقنين من العاديين درجة معيارية (ت) يمكن تحويل للدرجات الخام اليها بحيث تخطط النتائج على بطاقة صحيفة نفسية profile sheet وتتركز الأهمية الأولى على الدرجات التى تزيد عن ٧٠ والدرجة ٥٠ هى متوسط جماعة المرجع) ٠ وهذه الدرجة الفاصلة ، أى ٧٠ ، هى درجة مختارة على أساس قطعى اعتسافى ، أى ليس لها أساس أو تبرير نظرى ٠ وعلى أى حال نجد ان مفسرى نتائج الاختبار يميلون الى

الاهتمام بكل الارتفاعات ، أو القمم ، فى بطاقة الصحيفة النفسية ، سواء كانت هذه الارتفاعات تتخطى خط الـ ٧٠ درجة أو لا تتخطاه .

مقاييس الضبط : ومن العلامات المميزة والخاصة بمقياس مينسوتا المتعدد الأوجه استعماله لأربعة مقاييس تسمى مقاييس الصدق ، والصدق هنا معناه الضبط Control لأنها فى الواقع تمثل ضوابط للاهمال ، وسوء الفهم ، والتحليل ، ولتأثير انماط الاستجابات والاتجاهات نحو أخذ الاختبارات ، وهذه المقاييس هى :

١ ـ الأول ؟ ، وتمثل درجة هذا المقياس مجموع العناصر التى أجاب عنها المفحوص بالا أستطيع أن أجزم ، ومن المرغوب فيه أن تكون هذه الدرجة أقل ما يمكن .

٢ ـ الثانى ل L ، ويمثل الكذب ، ويعتمد على مجموعة من العناصر التى تجعل المفحوص يبدو فى شكل مقبول اجتماعيا ، ولكن ليس من المحتمل أن تكون الاجابة عنها صادقة فى الاتجاه المرغوب فيه . ومن امثلة هذا العنصر .

لا أقول الصدق دائما .

وعلى الرغم من أن الاجابة الصحيحة المعتادة تكون نعم ، الا أن الاجابة المقبولة اجتماعيا تكون لا ، ومن المفترض أن الشخص الذى يريد أن يظهر نفسه فى صورة مقبولة يحصل على درجة مرتفعة فى المقياس ل عن طريق تحريف استجاباته لعبارات المقياس ، وتدل الدرجات العالية على هذا المقياس على أن الاجابات لا يمكن الثقة بها ، ولكنها لا تدل على أن المفحوص يتعمد الكذب .

٣ ـ الثالث ف F ، وهو مقياس الخطأ ، ويتكون من العناصر التى تدر أن أجيبت من مجموعة التقنين فى نفس الاتجاه الذى أجابه فيها المفحوص ، وعلى الرغم من أن هذه العناصر لا يمكن أن تمثل نمطا معينا من الشذوذ النفسى فانها تعبر عن سلوك غير مرغوب فيه ، وعلى هذا فانه من غير المحتمل فى الواقع أن يعبر مفحوص واحد عن كل أو معظم هذه الأعراض

وترتفع الدرجة على هذا المقياس عندما لا يستطيع المفحوص أن يعطى اجابة مميزة لسبب من الأسباب كأن لا يستطيع القراءة أو الفهم بدرجة مناسبة ، أو الاهمال ، سواء عن قصد أم غير قصد ، كما قد يرجع ارتفاع الدرجة الى الخطأ فى التصحيح ، وبصفة عامة كلما زادت الدرجة على هذا المقياس زاد احتمال أن بعض العوامل تدخلت لتقلل من صدق الصحيفة النفسية (أو البروفيل) .

٤ - الرابع ك K ويمثل مقياس التصحيح ، وفى هذا المقياس استعمال لتركيبة أخرى من العناصر المختارة ، وتعطى درجة هذا المقياس دلالة على الاتجاه نحو أخذ الاختبار ، وهى متصلة بكلا المقياسين السابقين ل ، ف . ولكن يعتقد أنها أكثر خفاء منهما وتدل الدرجة العالية على هذا المقياس على اتجاه دفاعى defensiveness او على محاولة المخادعة فى الاتجاه الحسن « Fake good » او التحريف المقصود نحو الطرف السوى . كما تدل الدرجة المنخفضة على الصراحة المفرطة ، ونقد الذات ، او المحاولة المتعمدة للمخادعة فى الاتجاه السيء fake bad او الميل الى اظهار النفس بالمظهر غير السوى .

ومن المهم أن نشير الى أن استعمال مقاييس الضبط الأربعة هذه ليس مقننا تماما ولكنه يترك جزئيا لحكم مستعمل الاختبار ، وكذلك نشير الى أن هذه المقاييس تخضع باستمرار لاعادة التحديد والمراجعة فى ضوء البحوث الجديدة .

كانت تلك هى المقاييس الاكلينيكية التسعة الأساسية ثم المقاييس الأربعة الأخرى الخاصة بضبط الاختبار ، الا أنه منذ نشر الاختبار أنشىء عدد كبير جدا من المقاييس الجديدة بواسطة باحثين آخرين لم يشاركوا فى وضع المقاييس الأصلية ، من هذه المقاييس الجديدة مقياس عن الانطواء الاجتماعى (س ى) ، Social Introversion (Si) وقد أضيف الى المقاييس التسعة الأصلية وهو مضمن فى الاختبار ورقمه (صفر) . ويقيس هذا المقياس النزعة الى الانزواء والابتعاد عن الاتصال الاجتماعى بالآخرين .

وبهذا فان الاختبار فى صورته الحالية يعطى ١٤ درجة على الاربعة عشر مقياس ،نها التسعة الأصلية ، ومقياس الانطواء الاجتماعى ، ثم مقاييس الضبط الأربعة .

ومعايير المجموعة الضابطة الأصلية والتى تتكون من حوالى ٧٠٠ فردا مذكورة فى صورة درجات تائية (ت) . أو درجات معيارية بمتوسط ٥٠ وانحراف معيارى قدره ١٠ . وتستعمل هذه الدرجات المعيارية فى رسم الصحيفة النفسية ، وأى درجة تعادل ٧٠ أو تزيد عنها ، أى تقع فوق المتوسط بدرجتين انحراف معيارى ، ينظر اليها بصفة عامة على انها تميز انحرافا مرضيا ومما يجب ملاحظته ان المعنى الاكلينيكى لنفس الدرجة قد يختلف من مقياس الى مقياس اخر . فمثلا الدرجة ٧٥ على مقياس توهم المرض وعلى مقياس الفصام لا تدل على نفس الدرجة من حدة الانحراف .

اما تفسير الدرجات المنخفضة على الصحيفة النفسية للاختبار فليس على درجة كافية من الوضوح . اذ يعتقد الاكلينيكيون أن وقوع درجات اقل من ٥٠ كثيرا قد يكون له مغزى تشخيصى . ولكن لم توضع بعد تفسيرات منظمة لهذا .

على أى حال ، فان هناك قدرا من نتائج البحوث التى اجريت على هذا الاختبار يوضح أنه بصفة عامة كلما زاد عدد وحجم الدرجات المنحرفة على الاختبار ، زاد احتمال أن يكون الشخص مصابا باضطراب شديد .

ومن الواضح أن الاستعمال الرئيسى للاختبار المتعدد الأوجه هو التشخيص المتمايز أو الفارق differential diagnosis عندما يستعمل فى هذا المجال لابد أن يمارس مستعمله منتهى الحذر ولا يلجأ الى التفسير الحرفى للمقاييس الاكلينيكية فيه ، لأن الأمر اكثر تعقيدا من تلك الأسماء المعطاة للمقاييس الأصلية . فليس معنى الدرجة المرتفعة على مقياس الفصام مثلا أن الشخص مصاب بالفصام . فقد يحصل شخص عادى على نفس الدرجة . بينما قد يحصل المصابون بالفصام فعلا على درجات مرتفعة على مقاييس اخرى . ولكى يستبعد احتمال سوء التفسير للدرجات على المقياس الواحد فان ارقاما من صفر الى ٩ حلت محل أسماء المقاييس فى طبعات

أخيرة من الاختبار ٠ ولكى يسهل النظر فى نمط الدرجات أو الصحيفة النفسية وفحصها وضع نظام ترميز عددى للصحيفة النفسية ٠ وفى هذا النظـام يدل تتابع وترتيب أرقام المقاييس ، بنظـرة سريعة ، على النقط المرتفعـة والمنخفضة فى الصحيفة النفسية لشخص ما ٠ فمثلا الرمز ٩٤ ـ ٢ يدل على درجة مرتفعة على المقياس رقم ٤ (ب د) وعلى المقياس رقم ٩ (م ١) بترتيب متناقص ، أى أن الدرجة على المقياس الأول أعلى منها على المقياس الثانى ، كما يدل هذا الرمز على درجة منخفضة على المقياس رقم ٢ (د) ٠

والواقع أن أسماء المقاييس التى استبعدت كانت تستند الى تقسـيم اكلينيكى يقوم على أساس نظرى مشكوك فيه على الرغم من شيوعه وقت وضع الاختبار ، فالتقسيمات المستعملة والتى وضعت لها مقاييس الاختبار كانت محل شك ومراجعة فى علم النفس المرضى ، أما فى المجال العلمى الواقعى فان هذه التقسيمات لم تعد مقبولة ٠

ومن دواعى الحذر فى تفسير درجات المقاييس أن بعض هذه المقاييس لا تتمتع بثبات مرتفع ٠ وتعتبر هذه نقطة ضعف فى الاختبار ٠ لأن فاعلية أى تحليل للصحيفة النفسية تقل بسبب أخطاء الصدفة فى الدرجات التى تعتمد عليها فى هذا التحليل ٠ فاذا كانت درجات المقاييس المفردة غير ثابتة ومتداخلة أى تترابط فيما بينها ترابطا عاليا فان كثيرا من الاختلافات فيما بين الدرجات ربما تكون ناتجة عن الصدفة ٠

وتتراوح معاملات ثبات الاختبار الناتجة عن اعادة الاختبار ، بالنسبة الى الراشدين الأسوياء وغير الأسوياء بين ٥٠ر و ٩ر٠ وكانت الفترة بين الاختبار واعادته تتراوح بين بضعة أيام الى عام ٠ الا أن هناك دراسـة أجريت على طلاب الجامعة كانت معاملات الثبات فيها عموما أقل من السابقة، كما أظهرت الدراسة نفسها تنوعا كبيرا بين مقاييس الاختبار كانت تتراوح بين ٠٥ر٠لقياس ب ١ ، ٨١ ٠ لمقياس ب ت ٠

وثمة عيب آخر فى الاختبار يتصل بمدى تمثيل عينة التقنين ٠ فالدرجات المعيارية التى اشتقت منها تنظيمات الصحائف النفسية موضوعة فى ضـوء جماعة ضابطة تتكون من ٧٠٠ راشد هم الذين اختبروا فى التقنين الأصلى

للاختبار · وتبدو هذه العينة قاصرة جدا اذا قورنت بعينات التقنين الأخرى الأكثر شمولا والتى استعملت ، مثلا فى تقنين اختبارات القدرة وقد أكدت ابحاث كثيرة ضرورة أن تكون المعايير متنوعة تنوعا كبيرا فى جماعات أصل مختلفة ، ومن هذه الابحاث ما وضح أن متوسطات درجات طلبة الكليات كانت فوق ٥٠ بشكل مطرد على بعض المقاييس ، كما بينت دراسة أخرى أن ٣٩٪ من طلبة الكليات حصلوا على درجات فوق ٧٠ على مقيـــاس او اكثر من مقاييس الاختبار ·

ويتأثر الاختبار المتعدد الأوجه ، مثل سائر اختبارات الشـــخصية ، بالاختلافات الثقافية العامة بين المجتمعات بل وبالمستويات الثقافية الفرعية داخل المجتمع · وقد اظهرت دراسات كثيرة أجريت فى بلاد مختلفة ارتفاعا ملحوظا فى بعض مقاييس الاختبار عندما وضعت الصحائف النفسية على المعايير الأمريكية الأصلية · وان تفسير مثل هذه الاختلافات فى ثقافة المجتمع او ثقافاته الفرعية يتطلب معرفة بخصائص هذه الثقافات وظروفها ، فقـــد يكون الاختلاف فى درجات الاختبار المتعدد الأوجه راجعا الى اختلاف الأفراد فى تفسير عبارات الاختبار أو تعليمات الاختبار ، فمثلا قد يرجع الارتفاع فى بعض المقاييس عند بعض الجماعات الى اعتقادات متوارثة نحو التواضع وتقليل شأن الذات · وكذلك تتأثر الدرجة بالاختلافات الثقافية يتصل بالسلوك الذى يعتبر مقبولا اجتماعيا · الا أن الدرجات المرتفعة على مقاييس هـذا الاختبار ، بالنسبة الى بعض المجتمعات والجماعات ، قد تدل على وجود مشكلات انفعالية حقيقية يمكن ارجاعها الى أساليب تنشئة الطفـــل ، أو الصراعات فى الأدوار الاجتماعية ، أو غير ذلك من أســـاليب الاضطراب الانفعـــالى ·

اما عن الترجمة العربية لاختبار الشخصية المتعدد الأوجه فقد أعدها عطية هنا ، ومحمد عماد الدين اسماعيل ، ولويس كامل مليكه · ووضعت له معايير من البيئة المحلية ، وحسبت معاملات الثبات للمقاييس بطـــريقة للتجزئة النصفية للمقياس ، وبطريقة كيودر ـ ريتشاردسون · وقد استخدم الاختبار فى صورته العربية فى عدد من الدراسات (١٠ : ٣١٦ ـ ٣٢٤) ·

اختبارات مشتقة من اختبار الشخصية المتعدد الأوجه :

الى جانب حركة البحوث الضخمة التى حركها الاختبار المتعدد الأوجه فقد كان له فى ذاته خصوبة عالية من حيث اشتقاق مقاييس جديدة منه أو تأسيس اختبارات وتطويرها معتمدة عليه ·

فمن المقاييس الجديدة الكثيرة التى استخرجت منه عن طريق استعمال مفاتيح تصحيح جديدة نذكر ما يلى :

١ ـ مقياس التعصب (Pr.) ويهدف الى قياس العوامل السيكولوجية المرتبطة بالتعصب ضد جماعات اقلية ·

٢ ـ مقياس المكانة الاجتماعية ـ الاقتصادية (St) ويقيس بعض العوامل السيكولوجية المرتبطة بالمكانة الاجتماعية ـ الاقتصادية للشخص ·

٣ ـ مقياس الانحراف الجنسى ·

٤ ـ مقياس قوة الأناة ·

ويعتبر مقياس تايلور للقلق الظاهر (٥٩) الذى سبقت الاشارة اليه ، من المقاييس التى اعتمدت على اختبار الشخصية المتعدد الأوجه · وقد استعمل هذا المقياس استعمالا واسعا فى البحث وفى الممارسة الاكلينيكية · وقد انشىء أصلا للاستعمال فى التجارب لاختبار فروض معينة تتصل بتأثير الدافع على التعلم ، وتحقيقا لهذا اختيرت عبارات معينة من الاختبار المتعدد الأوجه على انها تدل على تعبير ظاهر عن القلق ، وكانت نتيجة هذا الاختبار ٥٠ عبارة ، زيد عليها عبارات اضافية لا تتصل بالقلق ومختارة من مقاييس الضبط الثلاثة ل ، ف ، ك · وقد ترجمه الى العربية الدكتوران مصطفى فهمى ومحمد احمد غالى ·

كذلك نتج عن الجهود التى بذلت لتعديل الاختبار المتعدد الأوجه بحيث يستعمل مع تلاميذ المرحلة الثانوية والطلبة الجدد فى الكليات أن أنشىء

اختبار منسوتا للارشـــاد النفسي Minnesota Counseling Inventory

ويتكون هذا الاختبار من ٤١٣ عبارة على المفحوص أن يجيب عنها بصواب أو خطأ ، ومعظم العبارات مأخوذة من اختبار الشخصية المتعددة الأوجه . ويعطى اختبار منسوتا للارشاد النفسي درجات في سبعة مجالات هي العلاقات الأسرية ، والعلاقات الاجتماعية . والثبات الانفعالي ، والمسايرة ، والتكيف مع الواقع ، والمزاج ، والقيادة . وعلى الرغم من أن مقاييس اختبار الارشاد النفسي تسمى باسم الطرف المقبول ، أى الثبات الانفعالي مثلا بدلا من الطرف الآخر وهو الاضطراب الانفعالي ، فانها تتشابه مع مقاييس الاختبار المتعدد الأوجه المسماة بأسمائها المرضية، فمثلا مقياس «المسايرة» في اختبار الارشاد النفسي يشبه الى حد كبير مقياس الانحراف السيكوباتي (Pd) في الاختبار المتعدد الأوجه ، وكذلك مقياس التكيف مع الواقعين يشبه كثيرا مقياس الفصام (Sc) ولهذا الاختبار مقياسان للضبط يشبهان مقاييس الضبط في الاختبار المتعدد الأوجه وتشير بعض الدراسات (في ٢٧ : ٤٩٦) الى أن مقاييس هذا الاختبار ذات قدرة ايجابية ولكنها ضعيفة على تمييز التلاميذ الذين يعرف أنهم غير متوافقين أسريا من غيرهم ممن يتمتعون بتوافق جيد .

وللاختبار ثبات باعادته بعد ثلاثة شهور يتراوح بين ٧٠ر٠ ، ٨٠ر٠ الا أن تفسير هذا الاختبار يحتاج الى حذر كبير . وقد نقل هذا الاختبـار الى اللغة العربية محمد عماد الدين اسماعيل وسيد عبد الحميد مرسى باسم مقياس الارشاد النفسي ، ولا تتوافر عنه بيانات سيكومترية في البيئة المصرية .

ومن الاختبارات التى اعتمدت على الاختبار المتعدد الأوجه اختبـار كاليفورنيا النفسي California Psychological Inventory فهو يشتق نصف عباراته من الاختبار المتعدد الأوجه ومن أمثلة مقاييسه : السيطرة ، النزعة الاجتماعية ضبط النفس ، الانجاز عن طريق المسايرة ، والانجاز عن طريق الاستقلال والمرونة ، وتعتبر انستازى (١٨ : ٤٤٨) هذا الاختبار من افضل اختبارات الشخصية المستعملة حاليا ، وقد خضع لدراسات وتحسينات عديدة أعطته هذه المكانة . ولهذا الاختبار صورة لتلاميذ الثانوى هى اختبار كاليفورنيا للشخصية California Personality Inventory نقلهـا الى اللغة العربية جابر عبد الحميد ويوسف الشيخ ، ولا تتوافر عنها بيانات سيكومترية في اللغة العربية . كما توجد صورة من هذا الاختبار لتلاميـذ

المرحلة الابتدائية نقلها الى اللغة العربية عطية محمود هنا باسم « اختبار الشخصية للأطفال » (*) .

نماذج أخرى ممثلة من اختبارات الشخصية المعتمدة على التقرير الذاتى

عرضنا بشيء من التفصيل اختبار الشخصية المتعـــددة الأوجه لأنه « النموذج الأمثل » لاختبارات الشخصية التى تعتمد على التقـــرير الذاتى للمفحوص ، الا ان الاختبارات التى تنتمى الى هذا النوع كثيرة ومتنوعة الى درجة يكاد يستحيل حصرها جميعا وخاصة فى هذا المجال المحدود ، لهذا نختار نماذج ممثلة من هذه الاختبارات التى لم تناقش فى الصفحات القليلة الماضية مراعين أن يكون لكل منها ميزته التى تجعله جـــديرا بهذا العرض الموجز .

قائمة مونى للمشكلات Mooney Problem Check list وقد وضعها روس مونى Ross L. Mooney وزملاؤه . ولها عدة صور من مراحل التعليم المتوسط حتى الجامعة والراشد . ويحدد فيها المفحوص مشكلاته فى احد عشر مجالا منها الأخلاق والدين ، النواحى المالية والمعيشية ، والتوافق للعمل المدرسى ، والعلاقات الاجتماعية ، وتشير الدرجة المرتفعة الى الحاجة الى الاشارد النفسى ، وتتخذ العناصر التى يحددها المفحوص أساسا للمناقشة الفردية والجماعية وقد استخدمت هذه القائمة فى عدة دراسات عربية .

اختبار التوافق لـــ بل The Bell Adjustment Inventory

منذ صدور هذا الاختبار فى عام ١٩٢٤ (٢٤) وهو يستعمل على نطاق واسع . والصورة الخاصة بالطالب مصممة للانتقاء السريع لطلاب المدارس الثانوية والكليات بهدف الارشاد النفسى ، وهذه الصـــورة تعطى

(*) أجرى عطية محمود هنا دراسة على « اختبار الشخصية للاطفال » استعرض فيها صدقه وثباته فى البيئة المصرية . راجع :

عطية محمود هنا اختبار الشخصية للاطفال وقيمته فى البحوث النفسية . المجلة الاجتماعية القومية مايو ١٩٦٥ .

درجات توافق في مجالات أربعة هي : الأسرة ، الصحة ، الناحية الاجتماعية، والناحية العاطفية · والصورة الأخرى خاصة بالكبار وهي اقل استعمالا ، وقد زيد عليها مجال خامس هو مجال التوافق المهني · وقد ترجم الصورة الخاصة بالطلاب محمد عثمان نجاتي · ولا تتوافر عنه بيانات سيكومترية في البيئة المصرية ·

مقياس البورت لدراسة السيطرة والخضوع
Allport A-S Reaction Study

يعتبر هذا المقياس من اقوى اختبارات الشخصية المبكرة ، ويوصف بانه يقيس السيطرة ascendance والخضوع submission لأنه يقيس نزعة الشخص الى أن يسود ويسيطر على زملائه ، او ان يساد ويسيطر عليه في علاقات المواجهة اليومية · ويبدأ كل عنصر بوصف موجز لموقف من مواقف الحياة اليومية العادية في اجتماع أو في مدرسة او سيارة عامة ، ويذكر أسلوبين او اربعة لمواجهة الموقف ، وعلى المفحوص أن يحدد اى الأساليب يعبر اكثر عن سلوكه · وتختلف الاستجابات في درجة السيطرة والخضوع التي تمثلها وتعطى لها أوزان مناسبة لهذه الدرجة في التصحيح · وفضلا عن شيوع هذا الاختبار فانه قد أثر في كثير من الاختبارات التي انشئت بعده ومن المهم ان نشير الى أن السيطرة كانت احدى السمات التي برزت بوضوح في دراسات التحليل العاملي ·

البروفيل الشخصي لجوردن Gordon Personal Profile

يعتبر هذا الاختبار ، وقائمة الشخصية للمؤلف نفسه ، من اختبارات الشخصية التي تعتمد على الاختيار الاجباري forced - choice للتغلب على ما يسميه المرغوبية الاجتماعية ، اى ميل المفحوص الى أن يبدو مقبولا اجتماعيا ، وكذلك للاقلال من الغموض وبعض الصعوبات الأخرى التي تصاحب عملية تصحيح اختبارات الشخصية ·

ويقيس البروفيل اربعة جوانب للشخصية هي : السيطرة ، والمسئولية ، والاتزان الانفعالي ، والاجتماعية (أو النزعة الاجتماعية) وهذه الجوانب

مستقلة نسبيا . والبروفيل مناسب للاستخدام مع طلاب المدارس الثانوية والجامعة والراشدين .

ويتكون البروفيل من ١٨ مجموعة من العبارات الوصفية تشمل كل مجموعة على أربع عبارات وتمثل كل عبارة احدى سمات الشخصية الأربع ، المشار اليها من قبل . وتتشابه جملتان من الأربع من حيث أن لها قيمة تفضيلية عالية . اما الجملتان الأرتان فتتساويان فى القيمة التفضيلية المنخفضة . ويطلب من المفحوص أن يضع علامة على جملة من الجمل الأربعة على أساس انها تشبهه الى حد بعيد وعلى جملة اخرى على اساس انها تشبهه الى اقل درجة وبهذا لا يستطيع المفحوص أن يستجيب استجابة قبول لجميع العبارات كما يحدث فى كثير من مقاييس التقرير الذاتى الأخرى أى أن البروفيل يعتبر اقل قابلية للتشويه من جانب أولئك الذين يحاولون أن يعطوا صورة طيبة عن أنفسهم .

وقد اقتبس البروفيل واعده بالعربية جابر عبد الحميد جابر ، وفؤاد أبو حطب ، ووضعا له كراسة تعليمات تتضمن وصفا له وتعليمات التصحيح كما وضعا له معايير مئوية ، وأجريت دراسات عن معاملات المقاييس الأربعة للبروفيل وصدقه ، كما استعمل فؤاد أبو حطب البروفيل فى دراسة (١٤) جاءت مدعمة لصدق التكوين الفرضى .

. قائمة لشخصية لجوردون Gordon Personal Inventory

وتعطى هذه القائمة قياسا سريعا ومناسبا لأربع من سمات الشخصية هى : الحرص ، والتفكير الأصيل ، العلاقات الشخصية ، الحيوية . ومن الممكن استخدامها فى المدارس الثانوية والجامعة ومع الراشدين . وتستخدم هذه القائمة ، مثل البروفيل الشخصى ، طريقة الاختيار الاجبارى ، وهى تتكون من ٢٠ مجموعة من العبارات الوصفية ، وتشتمل كل مجموعة منها على اربع عبارات ، وتمثل كل عبارتين يعتبرهما الأشخاص العاديون متساويتين فى قيمة التفضيل العالية ، وعبارتين يعتبرونهما متساويتين فى قيمة التفضيل المنخفضة . ويطلب من المفحوص أن يضع علامة أمام عبارة واحدة من العبارات الوصفية على أساس انها تنطبق عليه اكثر من الصفات الثلاث

الأخرى ، كما يضع علامة أخرى أمام عبارة أخرى باعتبارها أقل انطباقا عليه من العبارتين الاخرتين ، وتتميز القائمة بسهولة التطبيق والتصحيح والتفسير • وقد ترجمها الى العربية فؤاد أبو حطب وجابر عبد الحميد جابر ووضعا لها كراسة تعليمات تتضمن وصفا لها وتعليمات التطبيق والتصحيح كما وضعت لها معايير مئوية ، وحسبت لهـا معاملات الثبات والصـدق، واستعملها فؤاد أبو حطب فى الدراسة التى سبقت الاشارة اليها (١٤) •

• مقياس المسئولية الاجتماعية :

تـاليف سـيد أحمـد عثمـان ، وهـو مقيـاس لبعـد من الأبعـاد الاجتمـاعية ـ الأخـلاقية للشخصية على أسـاس تصور وتحليل نظرى للمسئولية الاجتماعية متمثلة فى عناصرها الثـلاثة وهى الاهتمام ، والفهم ، والمشاركة ، وفى المستويات المتدرجة فى كل عنصر من هذه العناصر، كما حددها وناقشها مصمم المقياس (٥) • ويتكون المقياس من صورتين الصورة ت لتلاميذ الثانوى والصورة ك لطلاب الجامعة والراشدين •

وتتكون الصورة ت من المقياس من ١١٥ عنصرا ، وهى عبارات تعكس الوانا من السلوك أو الآراء ، منها عبارات موجبة واخرى سالبة ويطلب من المفحوص أن يحدد اجابته وفق مقياس متدرج من أربع نقط •

وتسير الصورة ك من المقياس على النمط ذاته ، وهى مقتبسة ومختارة من الصورة الأولى الا انها تتكون من ٨٥ عنصرا ، منها ايضا عناصر موجبة واخرى سالبة •

وتتضمن كراسة تعليمات المقياس ، بصورتيه ، شرحا للتحليل النظرى للمسئولية الاجتماعية ، ثم بيانات عن وصف كل صورة من المقياس ، وعن صدق الصورة ت وثباتها ، واشارة الى الدراسات التى أجريت على هذه الصورة حتى الآن ، ثم طريقة التصحيح بالنسبة الى الصورتين ، وفى كراسة التعليمات ايضا بيان بامكانيات الاستعمالات المختلفة للمقياس فى صورتيه فى النواحى التربوية ، وفى دراسة الشخصية ، وفى العـلاج النفسى ، وفى الدراسات الاجتماعية •

والدراسات مستمرة فى أعداد معايير المقاييس واستكمال دراسات الثبات والصدق وخاصة للصورة ك وقد أجريت طوال العقد المنقضى من منذ ظهور المقياس عدة دراسات عليه أكدت فعاليته ٠ ويعتبر هذا المقياس أداة بارعة مرجوة تحتاج الى عمل طويل ممتد متنوع تزداد معه الأداة دقة وكفاءة ، كما يزداد معه التصور النظرى للمسئولية الاجتماعية اتساعا ووضوحا وعمقا ٠

• اختبارات كاتل في الشخصية :

تعتمد هذه الاختبارات علىنتائج البحوث العاملية التى اجراها ريموند كاتل R. Cattell ومعاونوه ، وقد يكون اشهرها استخبار العوامل الستة عشر فى الشخصية ٠

Sixteen Personality Factor Questionnaire (16 PF).

الذى يقيس ١٦ سمة ، كما أعدت قائمة أخرى للمراهقين (من ١٢ ـ ١٦ سنة) وتسمى Jr. — Sr. High Scool Personality Questionnaire

وأعدها باللغة العربية سيد محمد غنيم وعبد السلام عبد الغفار باسم « استفتاء الشخصية للمرحلة الاعدادية والثانوية » ، وتقيس ١٤ سمة ٠ وكذلك أعدت قائمة ثالثة للمستويات العمرية من ٨ سنوات الى ١٢ سنة وتسمى IPAT Children's Personality Questionnaire وأعدها باللغة العربية عبد السلام عبد الغفار وسيد محمد غنيم باسم استفتاء الشخصية للمرحلة الأولى » ٠ وقد حسبت المعايير التائية للاختبارين الاخرين على عينات مصرية كما حسبت ثبات استفتاء الشخصية للمرحلة الأولى الأولى بطريقة الصورة المتكافئة (ن = ١٩٢ طفلا) ، ولاستفتاء الشخصية للمرحلتين الاعدادية والثانوية بطريقة اعادة الاختبار (ن = ١٣٠ تلميذا)٠ الا انه لا تتوافر بيانات سيكومترية عن صدقها فى البيئة المصرية ٠

• قائمة أيزنك للشخصية :

هذا الاستخبار من اعداد عالم النفس الانجليزى ايزنك وتقيس الانبساط مقابل الانطواء ، والعصابية ٠ ونقلها الى اللغة العربية جابر

عبد الحميد ومحمد فخر الاسلام · ولا تتوافر لدينا معلومات سيكومترية عنها فى البيئة المصرية ·

· استبيان مستوى الطموح :

هـــذا الاســـتخبار من تاليف كاميليـــا عبـــد الفتـــاح ، ويتـــالف من ١٩ ســـؤالا مصـــنفة الى ٧ فئـــات هى : النظرة للحياة، الاتجاه نحو التفوق ، تحديد الأهداف والخطة ، والميل الى الكفاح ، وتحمل المسئولية والاعتماد على النفس ، والمثابرة والرضا ، بالوضـــع الحاضر والايمان بالحظ · وحسبت له المعايير المئينية (ن = ١٢٢ طالبة ، ١٩٤ طالبا) ، وقد ذكرت المؤلفة معاملات ثبات الاستخبار وصدقه وتمييزه بين الأسوياء والعصابيين (*) ·

(*) كاميليا عبد الفتاح : كراسة تعليمات استبيان مستوى الطموح ، النهضة المصرية ، ١٩٧١ ·

الفصل الرابع عشر

قياس الشخصية عن طريق الاحكام والملاحظات المنظمة

راينا فى قياس الشخصية عن طريق التقرير الذاتى أن الشخص يختار ويعطى تقديرات يرى أنها تعبر عن نفسه ، أما عند قياس الشخصية عن طريق الأحكام Judgments والملاحظات المنظمة systemic observations فان الذى يقوم بتقدير السلوك أو الأداء المميز للفرد أو المفحوص شخص أو أشخاص آخرون ، ويمكن أن نقسم هذا النوع من قياس الشخصية الى قسمين (٩) :

(أ) مقاييس التقدير وتقريرات القياس الاجتماعى ·

ratings and sociometric reports

(ب) الملاحظة المنظمة لعينات السلوك :

systematic observation of behavior samples

وسوف نتناول كلا من هذين القسمين بشىء من التفصيل ·

(أ) مقاييس التقدير وتقريرات القياس الاجتماعى ·

يمكن أن نصنف مقاييس التقدير وتقريرات القياس الاجتماعى الى نوعين على أساس من يقومون بهما ، النوع الأول يقوم به الرؤساء أو الكبار، بينما يقوم الأفراد أو الزملاء بالتقدير والحكم فى النوع الثانى ·

١ ـ أما عن تقديرات الرؤساء أو الكبار فنحن نلاحظ أن الرؤساء أو المدرسين مثلا يقومون باستمرار بتقديم نوع من الأوصاف لمرءوسيهم أو لمن هم فى كفالتهم ، الا أننا لا نستطيع أن نقارن هذه الأوصاف لأنها تختلف فى أسلوبها ، لهذا استعملت مقاييس التقدير حتى يتوافر قدر معقول من التشابه بين التقديرات ·

ويتكون مقياس التقدير من قائمة من السمات يجرى عليها التقدير ·

وقد يتكون المقياس من قائمة بسيطة من الصفات يتم الاختيار بينها ، أو قد يكون مقياسا متصلا يتضمن اقساما وصفيةمتعددة للاختيار منها٠ ٠ ويوضح المثال الثانى ، (فى ٢٧ : ٥٠٧) هذا النوع من مقاييس التقدير المتصلة :

١ ـ كيف تتاثر انت والآخرون بمظهره وسلوكه ؟	* يسعى اليه الآخرون * يحبه الآخرون * يحتمله الآخرون * يتجنبه الآخرون * لا مجال لملاحظته * يسعى الى اضافة أعمال جديدة الى نفسه * يكمل العمل الاضافى المقترح	وقائع مؤيدة
٢ ـ هل يحتاج الى تحريك أو انه يعمل من تلقاء نفسه ؟	* يعمل واجباته العادية تلقائيا * يحتاج الى تحريك احيانا * تحتاج الى تحريك كثير ليعمل واجباته العادية * لا مجال لملاحظته	وقائع مؤيدة

ولكن ـ نلاحظ أن مقاييس التقدير عن طريق الرؤساء تكمن فيها مصادر متعددة للخطأ منها مثلا ما يسميه كرونباك (٢٧) « خطأ الكرم » ويقصد به ميل من يقوم بالتقدير الى اعطاء تقديرات طيبة مؤيدة ، فاذا سئل مدرس مثلا أن يقدر تلميذا ما فيما يختص بالميل الى التعاون ، فانه سوف يقدر كل التلاميذ الذين لا يثيرون الشغب تقديرا عاليا جدا ، فى اقصى طرف المقياس بينما يهبط بتقدير القلة من التلاميذ مثيرى الشغب ٠ ويكون الارتفاع فى تقدير الغالبية رد فعل لتمييزهم عن مثيرى الشغب ٠ ومن اسباب خطأ الكرم أيضا ان الرئيس المقدر قد يحس ان مما يمس قدرته الرئاسية أن يعترف أن مرؤوسيه ليسوا اكفاء ، أو أن الرئيس قد يجد من الأيسر عليه أن يحكم احكاما طيبة على مرءوسيه اجمالا بدل ان يتملى ويطيل النظر ليميز بينهم تمييزا دقيقا ٠

ومن مصادر الخطأ أيضا فى هذا النوع من التقدير الغموض والابهام، فقد يفسر المقدر سؤالا أو مصطلحا تفسيرا مختلفـــا ، فقد يفسر شخص « القيادة » ، مثلا على أنها تركيز السلطة فى يد القائد ، والانفراد باتخاذ القرارات ، فعندما يقدر هذا الشخص مفحوصا تتوافر فيه هذه الصـــفات فانه يقدره تقديرا عاليا ومخالفا لتقدير شخص آخر للمفحوص ذاته ، لأن المقدر الأخير يفهم القيادة على أنها تشجيع للمرءوسين ومساعدة فى اتخـــاذ قرارات جماعية ٠

ويعتبر أسلوب الاستجابة Response style عند المقدر مصدرا آخر من مصادر الخطأ فى مقاييس التقدير بواسطة الرؤساء أو الكبار ، فمثلا نجد أن مقدار ما يندر أن يستعمل أطراف المقياس فى وصف المفحوصين ، بينما يميل شخص آخر الى الاسراف فى استعمال هذه الاطراف ، أى يميل الى تصنيف المفحوصين الى أقسام متميزة جدا ٠ ومن ثم نجد أن مثـل هذين المقدرين يختلفان عندما يقدران شخصا واحدا ٠

وما يؤثر فى تقديرات الرؤساء أيضا أن معلوماتهم عن المفحوصين قد تكون محدودة ، فقد يختلف تقدير مدرس التربية الفنية ومدرس اللغـــة العربية فى تقدير ردود فعل تلميذ معين للهبوط ، ويرجع الاختلاف الى أن كلا منهما يعرف التلميذ ويلاحظه فى عدد محدود من المواقف ، وعلى هذا تكون معرفته به محدودة وليست شاملة ٠

ومن مصادر الخطأ فى هذا النوع من التقدير « أثر الهالة » أو « وقع الهـــالة halo effect ويحدث هذا الأثر أو الوقع عندما يتأثر المقدر تأثيرا قويا بالانطباع العام عن المفحوص ٠ فيتكون عنده رأى عام أو تقدير عام عن قيمة الشخص يؤثر كثيرا فى تقديره لسمات هذا الشخص الخاصة ٠ وهذا النوع من الخطأ يجعل نمط السمات فى الفرد غامضا عند المقدر ٠

الا أنه من الممكن تلافى كثير من مصادر الخطأ هذه اذا راعينا :

١ ـ الاختيار السليم لمن يقومون بالتقدير ، من حيث قدرتهم ودرجة اندماجهم وتأثرهم شخصيا بتقديرهم للمفحوصين ، فقد سبقت الاشارة الى

أن المقدر قد يميل في تقديره حيز لصالح مرؤوسيه لأن في هذا تأييد لكفاءة رئاسته ٠

٢ ـ كلما زادت معرفة المقدر بالمفحوصين كانت تقديراته أكثر صدقا ٠

٣ ـ اذا استعملنا مقدرا بعينه أكثر من مرة فمن المفيد أن نحتفظ بتقديراته لمقارنتها والتعرف على أسلوب استجابته ، أو ما يمكن أن نسميه خطأه الثابت ٠

٤ ـ من الممكن زيادة ثبات مقاييس التقدير عن طريق الاعتماد على أكثر من مقدر واحد ٠ على أن يكونوا على معرفة كافية بالمفحوصين ٠

٥ ـ مما يزيد من قيمة مقاييس التقدير أن تعد بعناية ودقة ٠ وبصفة عامة نجد أن المقياس ذا الخمس نقط يعطي تمييزا أكثر دقة من المقياس أو القائمة التي تعتمد على اجابات نعم أو لا ٠ كما أن مقياس الخمس نقط ينبه الى أنواع متباينة من الاختلافات والانحرافات ٠ وكذلك يساعد الاختيار الاجباري الذي تكلمنا عنه في بعض اختبارات التقرير الذاتي ، على زيادة قيمة مقياس التقدير ٠ والهدف من الاختيار الاجباري هو فصل عمل المقدر فيما يتصل بوصف ما يعمله المفحوص عن تقويم هذا العمل ، أي أن المقدر يكون مسئولا عن وصف المفحوص ، بينما يقوم المسئول عن اتخاذ القرار بتقويم هذا النوع من مقاييس التقدير القائم على الاختيار الاجباري أكثر ملاءمة لقرارات المؤسسات فيما يتصل بالانتقاء والتصنيف ، ولكنها لا تناسب كثيرا عملية التوجيه التربوي أو الوصف التفصيلي للفرد ٠

أما عن صدق مقاييس تقدير الرؤساء فتشير الدراسات التي أجريت عنها الى ضرورة أخذ هذه المقاييس بحذر فيما يتصل بالحكم على الأفراد من جهة ، وباستعمالها للتحقق من صدق بعض مقاييس واختبارات الشخصية من جهة أخرى ٠ ومن أمثلة هذا أن طريقة تقدير الكبار استعملت في دراسة صدق مقياس المسئولية الاجتماعية ٠ الصورة ت ، الذي تناوله الفصل السابق ، حيث قام المدرسون بتقدير تلاميذ العينة على مقياس خاص بالمدرسين وكان معامل الارتباط بين تقدير المدرسين ودرجات التلاميذ على المقياس

٤٦ر٠ وقد اعتبر هذا العامل مقبولا ، وخاصة اذا أخـــذنا فى الاعتبار ان تقدير المدرس للتلميذ فى أى ناحية من نواحى السلوك تتدخل فيه عوامل كثيرة منها على سبيل المثال ، تحصيل التلميذ الدراسى ، ومستواه الاقتصـــادى الاجتماعى ، ودرجة جاذبية التلميذ الشخصية ٠ وتفضيل المدرس وتحيزه الشخصى وهكذا » ، (٢ : ٢٧) ٠

وقد أثبتت دراسة قام بها بيترز وكاميل (٥٢) أن ارتباط تقديرات الرؤساء لمدى معرفة العاملين ودرايتهم باعمالهم مع هذه المعرفة والدراية كما قيست باختبار خاص ، كان الارتباط بينهما ٣٥ر٠ بينما وصل هـــذا الارتباط فى بعض الأقسام ٥٥ر٠ ، كما أظهرت دراسات أخـــرى ارتباط التقديرات بطول المدة التى يعرف فيها المقدر المفحوصين ، وبمقدار حبه لهم (فى ٢٧ : ٥١٨) ٠

ولكن مقاييس تقدير الرؤساء والكبار ، على الرغم من حدودها وتعتبر مصدرا غنيا بالمعلومات التى لا يمكن الحصول عليها الا عن طريقها ٠

٢ _ والنوع الثانى من مقاييس التقدير هو تقدير الاقران peer ratings التى تعطى ملومات تعلو فى قيمتها على تقديرات الكبار أو الرؤساء ، وحتى عندما تتوافر تقديرات الرؤساء فان تقديرات الأقران والزملاء تعطى بيانات عن جانب مختلف من الشخصية ٠ فاذا كان رئيس واحد هو الذى يقـــوم بالتقدير فان قيام مجموعة من الأفراد او الزملاء بالتقدير سوف يعطى فرصة لعدد كبير من الأحكام مما يترتب عليه أن متوسط التقديرات بالنسبة الى سمة معينة من سمات الشخصية سوف تكون أكثر ثباتا من التقدير الذى يقوم به رئيس واحد ٠ ومن الممكن أن ننظر الى تقديرات الأفـــراد على أنها تعبير موضوعى عن سمعة الفرد ، فالسمعة تقوم الى حد ما على السلوك الفعلى ، الا أن نمط العلاقات الاجتماعية فى داخل الجماعة لا شك يؤدى الى وجود نوع من التحيز ٠

ومن الأساليب المستعملة فى تقدير الافراد ما يسمى أســـلوب تعيين الأســـماء nomination او الترشيح ، ويستعمل هذا الأسلوب كثيرا فى القياس الاجتماعى Sociometry ٠ وفى هذا الأسلوب يسأل كل عضو من

اعضاء الجماعة ان يسمى عددا محددا من الاشخاص الذين يتميزون بشكل واضح فى ناحية من النواحى ، مثل القيادة ، كما يسمى أيضا الاشخاص الذين تنقصهم هذه الخاصية ، وبالنسبة الى الأطفال فان عملية التسمية قد تتخذ صـــورة لعبة التخمين ، أو اختبار « خمن من » Guess Who الذى يصف الأدوار المختلفة التى يمكن ان يلعبها الاطفال ويسمى كل عضو فى الجماعة أسماء الأطفال الذين يعتقدون ان كل وصف يلائمهم • وترسم لكل طفل صفحة نفسية عن طريق عدد المرات التى اشير اليه فيها لكل وصف•

واذا كان من الممكن دراسة خصائص الفرد بطريقة خمن من ، فان الرسم الاجتماعى (السوسيوجرام Sociogram) يعطى فهما واستبصارا أعمق فى تمييز التجمعات الداخلية فى الجماعة ، وبيان البناء الهرمى للقيادة والمنعزلين ، اى هذا الرسم الاجتماعى يعتبر طريقة لدراسة البيئة الاجتماعية فى الجماعة (١٦ : ٤٤٣ وما بعدها) •

والذى أنشأ الرسم الاجتماعى وطوره مارينو Moreno (٢٩)• الا أن الأسلوب قد عدل بعدة طرق قللت من فاعليته فى سبيل الســهولة واليسر فى الاستعمال • وأفضل طريقة فى استعماله هى ان يسأل اعضاء الجماعة ان يختاروا رفقاء لهم فى نشاط معين • ولا بد ان تكون التعليمات متصــلة بانشطة حقيقية للجماعة ، كما يجب أن تكون الاختبارات حقيقية ويجب أن يعلم اعضاء الجماعة أن المعلومات التى يعطونها سوف تصان سريتها ، وأن البيانات المعطاة سوف تستخدم فيما اخبر اعضاء الجماعة انها ســوف تستخدم فيه ، مثلا لاختيار لجان ، أو جماعات عمل فرعيــة ، أو تنظيم الجلوس فى فصل دراسى وهكذا •

ويعتمد الرسم الاجتماعى الذى نحصل عليه على السؤال الذى توجهه، فان نمط الرسم الذى ينتج عن سؤال مجموعة من التلاميذ عن اختياراتهم لأقرانهم الذين يحبون أن يستذكروا معهم سوف يختلف عن نمط الرسم الناتج عن سؤال عن اختيارات من يحبون اصطحابهم فى رحلة •

ويشمل مصطلح التقدير بالقياس الاجتماعى Sociometric rating كل طرق تمييز العلاقات الاجتماعية بين اعضاء الجماعة • وليس هنـــاك

فصل قاطع بين تقدير الأفراد الوصفى ، وبين التقدير بالمقياس الاجتماعى ، الا اننا نجد النوع الأخير بصفة محدودة باسئلة عمن يحبهم المقدر اكثر أو يفضل العمل معهم ، ولهذا فهى تتصل باستجابة المقدر كما تتصل بشخصيته ·

ولتقديرات الأقران استعمالات متعددة منها · أن قائد الجماعة يستعملها فى التعرف على الأشخاص الذين يحتاجون الى عناية خاصة ، أو يستعملها فى اعادة تنظيم الجماعة حتى تؤدى وظيفتها بشكل افضل · كما أن هذه التقديرات يمكن أن تتخذ اساسا للانتقاء والتصنيف ، وهناك من يعتبرها انقى مقياس للقيادة ، أما بالنسبة الى المرشد النفسى ، فان تقديرات الأقران توضح خصائص الشخص التى تعوق قبول الجماعة له ، وخاصة عندما ينقص الشخص الاستبصار الكافى فى رأى الآخرين فيه ·

ويذهب كثير من الباحثين الى أن تقديرات الأقران اثبتت انها احدى وسائل التقدير التى يمكن أن يوثق بها الى درجة بعيدة لعدة أسباب هى :

١ _ أن عدد المقدرين كبير ، وقد أشرنا منذ قليل الى اهمية هذا فى زيادة ثبات التقديرات بصفة عامة ·

٢ _ ان أقران الشخص عادة ما يكونون اكثر قدرة من غيرهم على ملاحظة السلوك المميز له ، وهم من ثم أفضل فى الحكم على سمات معينة تتصل بالعلاقات الاجتماعية من المدرسين أو المشرفين او أى ملاحظين خارجين ·

٣ _ ان أراء أعضاء الجماعة الآخرين ، صوابا كانت ام خطا ، تؤثر فى سلوكهم وتصرفاتهم وبالتالى فان هذه الآراء تحدد جزئيا طبيعة تفاعلات الشخص مع الجماعة · فمثلا ، اذا كان أعضاء جماعة يرون ان أحد أعضاء جماعتهم ، وليكن اسمه عليا مثلا ، رجل يمكن الاعتماد عليه والثقة به ، فان هذا الرأى يوجه ويحدد نوع تعاملهم مع على ، وغالبا ما يحدد هذا العامل نوع السلوك من جهة على كما يحدد نوع التفاعل منجهة الجماعة فالجماعة تعامله على انه ثقة ، وهو يسلك تجاه الجماعة وفق توقعاته منه

(٥ : ٩٨ - ١١٠) ولا شك أن هذا يعطي تقديرات الأقران صفة أقرب الى الواقعية ·

(ب) الملاحظة المنظمة لعينات السلوك

من الممكن أن تعطينا الملاحظات المنظمة وصفا أكثر دقة للسلوك أو الأداء المميز اذا قورنت بأساليب التقرير الذاتي أو الأحكام والتقديرات التي يصدرها الآخرون عن الشخص ، ذلك لأنه في التقرير الذاتي من المحتمل أن يتدخل التذكر الانتقائي ، بل انه يحدث أيضا عندما تستعمل أساليب التقدير والحكم المختلفة . وكذلك نجد أن الملاحظة في هذه الأساليب الأخيرة ليست قائمة على الاختيار المنظم للمواقف التي يلاحظ فيها السلوك ، بل انها تقوم على ملاحظة المواقف كيفما اتفق وقوعها ، وفي هذه الأساليب أيضا لايلاحظ الشخص في كل المواقف · لهذا قلنا ان الملاحظة المنظمة تعطي وصفا أدق للسلوك المميز ·

ومن السهل أن نلاحظ سلوك مفحوص ما في ظروفه العادية · وهذا ما يسمى بالملاحظة الميدانية · هذا النوع من الملاحظة يكون أكثر ملاءمة في بعض الدراسات من الملاحظات المقننة ، فكثير من الباحثين يرى أنه من المستحيل أن نفهم الشخصية الا اذا لاحظنا المفحوص وهو يستجيب الى الظروف التي لها مغزى بالنسبة اليه ، لأن المثيرات المختلفة لها معان ومغاز مختلفة بالنسبة الى الأفراد المختلفين · وعلى هذا فهم يرون أن المواقف المقننة حيث تلاحظ ردود أفعال الأشخاص لموقف بعينه ، تبرز أنماط السلوك الهامة كما تبرزها الظروف العادية ، غير المتشابهة ، التي يعيش فيها الفرد · الا أن المشكلة هنا تكمن في كيف نرى قدرا كافيا من سلوك الفرد ، والحصول على تسجيلات لملاحظات السلوك يمكن الاعتماد عليها · أي أن المشكلة هنا ذات شقين : الأول ، عينة السلوك وما يرتبط بها من صعوبات ، والثاني ، عيوب أو أخطاء من يقوم بالتسجيل · ولنتكلم عن كل منهما بشيء من التفصيل ·

اختيار عينة السلوك :

لكي نعرف السلوك أو الأداء المميز لشخص ما لابد أن نعرف ما يميز تصرفه في موقف بعينه، ولكن المواقف تتغير من يوم الى يوم ومن لحظة الى

أخرى ، والطريقة الوحيدة لنطمئن الى أن ما نلاحظه هو السلوك المميز هى أن ندرس هذا الشخص فى مواقف ومناسبات عديدة ، وهو أمر مجهد ومكلف . فالواقع أنه لابد أن نصل الى نوع من التوفيق بين الاختيار الدقيق لعينة السلوك والاقتصاد فى الانفاق والجهد .

وانه من العسير جدا أن نقوم باستنتاجات عن فروق بين الأفراد من ملاحظة سلوكهم فى مواقف معينة ، لأنه لا يمكن ملاحظة فردين فى موقف واحد بعينه ، وحتى اذا كان الموقف واحد فى ظاهره فان الظروف السابقة الخاصة بكل منهما تجعلهما يسلكان سلوكا مختلفا . ولهذا فان أقصى مايمكن أن تقدمه لنا الملاحظات المقارنة هو بيان كيف يسلك اشخاص مختلفون فى ظروفهم الحالية ، ولكن هذه الملاحظات لا تضمن لنا أن تستمر هذه الفروق عندما تتغير طبيعة هذه الظروف .

ومن أفضل الأساليب التى تؤدى الى المقارنة الدقيقة ما يسمى بالعينات الموقوتة time sampling وفى هذا الأسلوب يخطط مسبقا جدول أو قائمة للملاحظات ويصمم هذا الجدول بشكل عشوائى يضمن أن يلاحظ كل مفحوص فى ظروف متشابهة مع الآخرين بحيث تمكن المقارنة . فمثلا عند دراسة الاتصالات الاجتماعية لأطفال ما قبل المدرسة يوضع جدول الملاحظات لمدة دقيقة واحدة ، وبعد ملاحظة الطفل لمدة دقيقة يكتب الملاحظ تقريرا كاملا عما لاحظه من تفاعلات الطفل الاجتماعية . ويلاحظ الأطفال وفق نظام مرسوم ويعدل من يوم الى يوم ، وفى خلال الدراسة كان كل طفل يلاحظ عدد مرات متساوية فى خلال الخمس دقائق الأولى من ساعات اللعب الحر ، ثم الخمس دقائق الثانية وهكذا . وللعينات الموقوتة قصيرة الأجل والموزعة توزيعا حسنا عدة مزايا ، فالصورة المتراكمة من الملاحظات فيها تكون اكثر تمثيلا للسلوك من نفس القدر من المعلومات التى يحصل عليها من ملاحظات من مرات اقل وذات فترات طويلة ، كما أن أخطاء الذاكرة تقل هنا لأن الملاحظ يقوم بالتسجيل أثناء وبعد الملاحظة مباشرة . والعينات الموقوتة اكثر مناسبة لتسجيل حقائق معينة يمكن أن تصاغ فى صورة رقمية ، مثل عدد مرات الاتصال الاجتماعى بالآخرين .

ومن الممكن الحصول على عينة اكثر اتساعا بأسلوب « السجل اليومى»

day record الذى أنشأه باركر وزملاؤه (٢١) • كان هدفهم هو دراسة نمط حياة الطفل ومع الاهتمام الخاص بالمجالات التى يتحرك فيها •

وللسجل اليومى مزايا أن يعطى صور عن تتابع النشـــاط ، ولكن له عيوبه ، فاذا كان الملاحظ معروفا ومكشوفا تأثر الســـلوك الذى يلاحظــه ، والاعتماد على يوم كامل بعينه لا يكفى للحصول على معلومات مميزة كاملة عن المفحوص •

• ومن المهم أن نلاحظ بالنسبة الى عينة السلوك الملاحظة عن طريق العينة الموقوتة ما يلى :

١ ـ أن الاستجابات التى يعدها الملاحظ متشابهة أو واحدة قد يكون لها فى الحقيقة معان مختلفة • وأن المواقف التى تبدو له متشابهة قد تستدعى استجابات متباينة •

٢ ـ ان الاستنتاجات المشتقة من موقف واحد ، حتى ولو كانت على أساس تراكم ملاحظات ، تكون ، أى الاستنتاجات صحيحة بالنسبة الى هذا الموقف فقط •

٣ ـ أن الملاحظات المطلوبة والمناسبة للحصـــول على معلومات ثابتة ويمكن الثقة بها ، يعتمد هذا العدد على المشكلة ، الا أنه بصفة عامة نجد أن عددا كبيرا من الملاحظات القصيرة افضل من عدد قليل من الملاحظات التى تعتمد على فترات طويلة •

عيوب من يقوم بالملاحظة :

الواقع أن بعض عيـــوب الملاحظين قد يرجع الى طبيعـــة المواقف الاجتماعية ذاتها التى تكون موضوع ملاحظة حيث تحـــول درجة تعقدها الكبيرة دون التقرير الشامل للمستغرق لها • أما العيوب الأخرى فترجع الى الملاحظ نفسه •

واذا كانت أخطاء الملاحظ مجرد حذف أو اغفال عشـــوائى لبعض ما

يلاحظ فليس لهذا وزن كبير ، واما ما يجب الاهتمام به فهو الأخطاء المنظمة ، كأن يبالغ باستمرار وبصفة مطردة فى تأكيد بعض انماط الأحداث ، كما يغفل الاشارة الى بعضها الآخر ·

عندما يلاحظ مراقبون موقفا واحدا فانهم يعطون تقريرات متباينة بدرجة كبيرة ، ذلك لأن كل ملاحظ عنده حساسية لأنماط معينة من السلوك اكثر من غيرها والوضع الأمثل هو ان يبنى الملاحظ انطباعه على كل تصرف يمكن أن يكشف أو يعكس خصائص اساسية ، ولكن عندما يركز من يقوم بالملاحظة والمراقبة على شيء بعينه فانه بالضرورة لابد أن يغفل ملاحظة اشياء أخرى ·

وثمة عيب آخر وهو أن الملاحظين يفسرون ما يرون ، فاذا قام الملاحظون بتسجيل الحقائق الموضوعية ، فقد يصل آخرون الى تفسيرات مختلفة تماما، ولكن الناس يميلون الى تفسير ما يرون ، وعندما يفسرون فانهم يميلون الى تجاهل الحقائق التى لا تتفق مع تفسيـــــيرهم بل انهم قد يخترعون حقائق لتكمل تفســـيرهم ·

ومما يقلل من آثار هذه العيوب أن يتم التسجيل بشكل منظم ، وان يكون كلما امكن ، فى صورة وحدات سلوكية يمكن عدها ، فمثلا من الممكن ان نصف مدى انتظام عمال مصنع فى عملهم عن طريق التســـجيل الزمنى للحظات التىيكونون فيها فىعمل فعلى، واللحظات التىينشغلون فيها عنالعمل للحصول على الأدوات ، أو زيارات زملائهم ، أو يتعرضــون فيها لمؤثرات التشتيت· وحتى اذا كان السلوك الملاحظ على درجة كبيرة من التنوع بحيث يصعب وضعه فى صورة وحدات سلوكية فمن الممكن ان تحدد اقسام السلوك وتعد قبل ويكون عمل الملاحظ هو أن يصنف السلوك الذى يلاحظه وفق هذا التقسيم ، ويضع علامة فى القسم المناسب لكل اسلوب يلاحظه كما هو الحال فى اقسام بيلز Bales (راجع ٦) ·

وعلى الرغم من فائدة التسجيل والعد الموضوعى لوحدات الســـلوك وملائمته للبحوث العلمية فان فائدته محدودة فى التوجيه والارشاد النفسى·

اما السجلات القصصية او الوقائعية anecdotal records فانها ·

وان كانت تنقصها دقة الموضوعية ، فانها تعطى وصفا حيا للشخص ، لأن الملاحظ فيها يوجه اهتمامه الى أى سلوك يبدو جديراً بالملاحظة بدل أن يركز على سمات معينة للمفحوصين كلهم الا ان الملاحظ فى السجل القصصى يصف بدقة ما يلاحظه فاصلا بين الوقائع وتفسيراتها · وعندما تتراكم هذه السجلات القصصية فانها تعطى صورة اكثر غنى من أى وسيلة أخرى على هذه الدرجة من البساطة · ومسئولية الملاحظ الكبرى هنا تتركز فى اختيار الوقائع التى تستحق التقرير · وفى أن يتحرى الموضوعية بقدر الامكان ·

الفصل الخامس عشر

قياس الشخصية عن طريق الأداء والأساليب الاسقاطية

على الرغم من أن اختبارات الأداء والأساليب الاسقاطية تستعمل فى قياس الشخصية منذ أكثر من أربعين عاما فانها لم تصل الى درجة النضج التي وصلت اليها الاختبارات والأساليب والطرق التى تكلمنا عنها حتى الآن. لقد كان الحلم الأكبر للمشتغلين بقياس الشخصية أن يصلوا الى وسائل قياس · تمثل السلوك مباشرة وليس انطباعات متميزة أو تقريرات ذاتية غير دقيقة ، ووسائل قياس تعطى نتائج كمية ، وتسمح بالمقارنة المباشرة بين الأفراد فى المواقف ذاتها · وكانت اختبارات الأداء والوسائل الاسقاطية ـ وليس بينهما تمييز قاطع كما يرى كرونباك (٢٧ : ٥٢٩) ـ كانت ضمن هـذه الوسائل ، وان كانت لم تحقق ، كما سبق ، نضجا يعادل ما وصلت اليه اساليب قياس الشخصية الأخرى ·

واختبارات الأداء التى تهدف الى قياس الشخصية تصمم بحيث تؤدى الى ظهور دلائل ومؤشرات وعينات من السلوك المميز حقا للشخص ، ويمكن أن نحدد الملامح العامة لمثل هذه الاختبارات فيما يلى :

١ ـ أن يكون الموقف المثير موحدا بقدر الامكان بالنسبة الى المفحوصين كلهم ·

٢ ـ يصمم الموقف بحيث يسمح بظهور تنوع فى أنماط السلوك التى يهدف المختبر الى ملاحظتها ·

٣ ـ يوجه المفحوص الى الاعتقاد أن خاصة معينة هى التى تقاس بينما يقوم الملاحظ بملاحظة خاصة أخرى أو جانب اخر من السلوك أو الأداء ·

٤ ـ يسجل الملاحظ بدقة طريقة المفحوص في الأداء وليس كمية السلوك المؤدى ·

والمزية الكبرى للملاحظة فى اختبارات الأداء انها تكشف عن خصائص لا تظهر الا قليلا جدا فى انشطة الشخص العادية من هذه الخصائص مثلا الشجاعة ، ورد الفعل للحبوط ، وعدم الامانة . ما ان من مزايا هذا النوع من الملاحظة فى مواقف الأداء ان رغبة المفحوص فى ان يعطى انطباعا حسنا لا تؤثر فى مدى صدق الاختبار ، لأن الشخص بسبب هذه الرغبة سوف يكشف عن خصائصه بدرجة أكبر مما لو لم تكن عنده هذه الرغبة ، وعلينا ان نأخذ فى الاعتبار هذا الدافع عندما نفسر نتائج ملاحظاتنا . ومن مزايا اختبارات الأداء أيضا انها تسمح الى درجة بعيدة بمقارنة الأشخاص فى تعرضهم لظروف واقعية موحدة ومتشابهة .

وتتراوح المواقف التى تستخدم الأداء بين المواقف الكاملة والمحددة فى بنيتها highly structured والمواقف التى لا تتميز ببنية واضحة محددة Totally unstructured . ويقال ان موقفا ما ذا بنية كاملة ومحددة اذا كان له معنى واحد معين لجميع المفحوصين . بينما يكون الموقف الذى لا تتوفر فيه بنية واضحة محددة متضمنا لعدد قليل جدا من الاشارات، او ليس فيه سوى نمط أو نظام ضئيل جدا الى درجة تجعل من الممكن للمفحوص أن يعطيه أى معنى يشاء . ومن أبسط امثلة المواقف غير محددة البنية الصوت الغريب الذي ينبعث من مكان فى المنزل مغلقا بستائر الظلام سابحا على أمواج السكينة فى وسط الليل .. تسمعه فلا تعرف بالتحديد ما هو ؟ مصدره ؟ يعطيه كل من يسمعه تفسيرا خاصا ، هذا التفسير الخاص يتأثر الى درجة بعيدة بمخاوفنا ، شعورية كانت أم لاشعورية ، وباهتماماتنا ومعرفتنا ، أى أنه فى الموقف محدد البنية يعرف الشخص بالضبط ما يتوقع منه ، وكيف يقوم به ، أما فى الموقف غير محدد البنية ، فان الشخص يوجه نفسه ويقودها فى الاستجابة له ، وكلما زاد غموض الموقف ، اى زادت بنيته عدم تحديد زادت الفرصة لتعطي الشخص تفسيرا او اداء خاصة مميزا له .

وتعتبر المواقف محددة البنية اكثر مناسبة لقياس القدرة ، او قياس اقصى الأداء ، لأنها تجبر كل مفحوص على أن يحاول أمرا بعينه ، وبينما نجد الأساليب غير محددة البنية ، او الأساليب الاسقاطية ، تستعمل ، بعكس مقاييس القدرة ، مثيرات لا تتميز غالبا باية بنية محددة .

وفيما يلى نتناول نوعين من اختبارات الأداء لقياس الشخصية ، النوع الأول هو اختبارات المواقف ذات البنية المحددة ، والنوع الآخر هو اختبارات المواقف ذات البنية الأقل تحديدا او الاختبارات الاسقاطية ٠

اولا : اختبارات الأداء ذات المواقف محددة البنية :

تدخل ضمن هذا النوع من الاختبـارات ٠ اختبارات الخلق والاصرار العمل ، كذلك الاختبارات التى تقيس الأساليب الادراكية والمعرفية ٠
perceptual and cognitive styles

لقد كان استعلام تربية الخلق The Character Education Inquiry

الذى وضعه هارتشون وماى Hartshorne and May من أكثر المجهودات التى بذلت لتقويم الشخصية عن طريق اساليب كمية وموضوعية خالصة ، وكان اعتقادهما أن سمات الخلق يمكن أن تقاس بدقة عن طريق تعريض الشخص للاغراء فى مواقف يظن انه يمكن فيها ان يخالف المعايير والقواعد الأخلاقية دون ان يكتشف أمره ٠ وكانت السـمات التى درسـها هذان الباحثان تتضمن ، الصدق ، والأمانة فيما يتصل بالنقود ، والاصرار والدأب فى العمل ، والتعاون ، الكرم ، (راجع ٢٢ ، ٢٣ ، ٢٤)٠

اما اختبارات الأفراد فى الأساليب والطـرق التى يتبعونها فى ادراك وتناول المشكلات ٠ والواقع أن الدارس الذى يقوم بملاحظة اى سلوك اثناء حل مشكلة يمكن ان يميز بسرعة الاختلافات فى الطرق التى يتبعها المفحوصون فى مواجهة المشكلات وتخطى العقبات ٠ وتعتبر الاختبارات الفردية ، مثل اختبار وكسلر الذى سبق تناوله ، فرصة لملاحظة هذه الأساليب او العادات المختلفة بين الأفراد ٠ والمعلومات التى نجمعها عن هذه الأساليب الخاصة تساعد فى التنبؤ بنوع المشكلات التى يستطيع الشـخص ان يتعامل معها بنجاح ، واى الاخطاء يمكن ان يقع فيها أكثر ، وقد تفتح امامنا طريقا لعلاج مناسب يؤدى الى تحسين الأداء فى مواجهة المشكلات ٠

وتهتم اختبارات العمليات المعرفية اساسا بالطـريقة التى ينظم بها

الشخص المعلومات التى ترد اليه ، هذا التنظيم Organising او اعطاء نمط patterning كان يمثل الاهتمام الرئيسى عند اصحاب نظرية الجشتالت ، ومعظم الاختبارات التى وضعت في هذا المجال وضعها باحثون ينتمون الى مدرسة الجشتالت · وقد استعمل هؤلاء الباحثون مثيرات بسيطة لاختبار الجهاز الادراكى كله معتقدين ان المخ جهاز لتحويل المعلومات وانه عن طريق مثيرات او اختبارات ادراكية بسيطة يمكن ان نتعرف على خصائصه المميزة· من هذه المثيرات الادراكية البسيطة اندماج الخفق flicker fusion والحركة الظاهرة apparant movement فعندما ينظر الشخص الى ضوء من خلال جهاز دوار يحجب الضوء ثم يكشفه ، فانه يدرك الخفق عندما يكون الدوران ذا معدلات بطيئة ، وكلما زاد معدل الدوران تدريجيا فان الشخص يستطيع ان يقرر النقطة التى يختفى عندها ادراكه للخفق ، وقد وجدت فروق فردية واسعة فى هذه الناحية ، ناك من الدلائل ما يشير الى أن نقطة الاندماج هذه تدل على قدرة الجهاز صبى على تسجيل تفاصيل المثيرات الواردة ، وقد استعملت مقاييس ندماج الخفق بنجاح فى تشخيص اصابات المخ ·

اما عن الحركة الظاهرة فالمقصود بها انه عندما يعرض ضوءان جنبا الى جنب مع اطفاء أحدهما وانارة الآخر · تتابع سريع فانه يظهر أن الضوء يتحرك قافزا جيئة وذهابا ، وفى الحركة الظاهرة هذه ، وتسمى أيضا ظاهرة فاى ، نجد أن المخ يقوم بتكامل المثيرات فى نمط · وقد درس كلين وشيليزنجر (٣٨) أثر تغيير الفترات بين الاضاءتين ، ووجدا انه كلما قصرت هذه الفترة كانت هناك عتبة فارقة لبدء ظهور الحركة الظاهرة وعتبة فارقة ثانية تختفى عندها ، أى يرى الشخص بعدها ضوءين ثابتين · وقد كان الفاصل بين هاتين العتبتين ، أى مدى الفترات التى تسمح بظهور انطباع الحركة ، اكثر اتساعا عند بعض الأفراد عنه عند غيرهم ·

ومن اختبارات الأساليب المعرفية ما يهدف الى قياس المرونة العقلية والجمود أو التصلب العقلى · فمما يلاحظ كثيرا أن الذين يفشلون فى حل المشكلات يتميزون بالتشتت والوقوف عند الأفكار الخاطئة ، فهم مثلا يكررون الدخول فى ذات الممر المسدود فى المتاهة · والواقع أن المواجهة السليمة للمشكلة تتطلب اعادة تنظيم المجالات الادراكية على اساس المعلومات

او المتطلبات الجديدة · وقد انشئت اختبارات لقياس درجة المرونة الادراكية والعقلية ، نشير منها الى اختبار التاهب Einstellung Test (او اختبار التوجيـــه orientation او اختبـــار الاتجـــاه العقلى mental set) الذى وضعه لوشينز (٤٤) وهو يعرف باسم اختبار اناء الماء · ويستعمل مشكلات توزيع كميات من المياه فى انية محددة الاتساع · ويجرى الاختبار هكذا مثلا ، اذا كان لديك اناء سعته ٧ ارباع جالون وآخر سعته ٤ ارباع جالون فكيف تحصل بالتحديد على ١٠ ارباع جالون ؟ » ، وتمشيا مع منطق الاختبار يسمى الاناء ٧ (ا) والاناء ٤ (ب) ويكون الحل هو : املأ ب تاركا ثلاثة ارباع جالونات فى ا ، افرغ ب ، املأ ب من ا ، ثم املأ ا · وعلى هذا يحصل على ثلاثة ارباع جالونات حسب القاعدة ا ـ ب ، وتحصل على الـ ٧ ارباع جالونات الأخرى من ا ، اى أن (ا ـ ب) + ا = ١٠ · وتكون سلسلة المشكلات كما يلى :

	الآنية			المطلوب	
	ا	ب	جـ		
مثال	٧	٤	ـ	١٠	٢ ا ـ ب
١ ـ	٢١	١٢٧	٣	١٠٠	ب ـ ا ـ ٢ جـ
٢ ـ	١٤	١٦٣	٢٥	٩٩	ب ـ ا ـ ٢ جـ
٣ ـ	١٨	٤٣	١٠	٢١	ب ـ ا ـ ٢ جـ
٤ ـ	٩	٤٢	٦	٢١	ب ـ ا ـ ٢ جـ
٥ ـ	٢٠	٥٩	٤	٣١	ب ـ ا ـ ٢ جـ
٦ ـ	٢٣	٤٩	٣	٢٠	ب ـ ا ـ ٢ جـ او ا ـ جـ اختبار
٧ ـ	١٥	٣٩	٣	١٨	ب ـ ا ـ ٢ جـ او ا + جـ اختبار
٨ ـ	٢٨	٧٦	٣	٢٥	ا ـ جـ ا انطفاء
٩ ـ	١٨	٤٨	٤	٢٢	ب ـ ا ـ ٢ جـ او ا + جـ اختبار
١٠ ـ	١٤	٣٦	٨	٦	ب ـ ا ـ ٢ جـ او ا ـ جـ اختبار

ومن الممكن مساعدة المفحوص فى حل المشـــــكلات الأولى · وتؤدى المشكلات من ١ حتى ٥ التى تحل عن طريق تطبيق قاعدة معينة الى تكوين « اتجاه عقلى او تأهب لاستعمال هذه القاعدة او المعادلة ، بعد ذلك تقدم

مشكلات الاختبار critical والانطفاء extinction . وفى مشكلة الاختبار كما فى المشكلة رقم ٦ مثلا يمكن ان تؤدى القاعدة او المعادلة الى الحل ، ولكن هناك طريقة أسهل للحصول على الاجابة . اما فى مشكلة الانطفاء كما فى رقم ٨ مثلا . فان التأهب ، اى استعمال القاعدة لا يؤدى الى الحل ولكن من الممكن بالنسبة الى المفحوص الذى يتمتع بمرونة أى يكتشف قاعدة اكثر بساطة ولكى يحصل المفحوص على درجة مرتفعة على اختبار التأهب هذا فان عليه ان يتنبه للمشكلة المباشرة أمامه ، مستبعدا أى ذكريات عن الحلول السابقة . ومن المحتمل يكون السلوك غير المرن ، المتصلب ، الجامد دليلا على عدم القدرة على الفصل بين المعلومات المتعارضة .

ومن الأمثلة الأخرى لاختبارات المرونة الادراكية اختبارات الأشكال المضمنة Embedded Figurs test حيث يقدم فيها شكل هندسى غريب ثم يطلب من المفحوص استخراجه من مجال أكثر تعقيدا ، وتكون الدرجة هى الوقت الذى يستغرقه المفحوص فى حل مثل هذه المشكلات . وقد أعده للاستخدام فى اللغة العربية سليمان الخضرى الشيخ وانور محمد الشرقاوى.

ثانيا : اختبارات المواقف غير محدد البنية

ينقسم هذا النوع من الاختبارات الى ثلاثة اقسام :

الأول ، اختبارات حل المشكلات ، والثانى ، الاختبارات الادراكية غير محددة البنية ، والأخير الاختبارات اللفظية غير محددة البنية . ولنتكلم عن كل قسم من هذه الأقسام .

(١) اختبارات حل المشكلات : من الممكن ان نلاحظ فى اختبار مثل اختبار تصميم المكعبات وبنائها الطريقة التى يتناول بها المفحوص المشكلة او ما يسمى بطريقة المواجهة method of attack كما يمكن ان نلاحظ رد الفعل للحبوط response to frustration عند هذا المفحوص كلما واجهته صعوبة او فشل فى الوصول إلى حل لما امامه من مشكلات . الا أننا نستطيع أن نحصل على معلومات افضل اذا عدلنا الاختبار ، واذا حددنا بالضبط ما نريد ملاحظته .

وقد حدد جولدنر (٣٠) Goldner طريقة المواجهة فى ضـــــوء متغيرين اكثر تحديدا هما :اتجاه الكل ـ الجزء whole-part approach والتصلب rigidity وقد استعمل ستة اختبارات منها : اختبار مكعبات معدل ، واختبارات تكوين كلمات من عدد معين من الحـــــروف ، واختبار الرورشاخ ٠ واختبـــار الوظيفة ، والذى يسال فيه المفحـــوص أن يعطى استعمالات او وظائف مختلفة لأشياء معينة مثل الصندوق أو المكنســـة او الورق ، وقد انشا جولدنر قواعد تصحيح لكل اختبار ، فمثلا بالنسبة الى اختبار الوظيفة تعطى درجة الكل ـ الجزء حسب ما اذا كانت الاجابة تشير الى استعمال الشيء كله ، مثلا وضع أشياء فى الصندوق ، أو كانت الاستجابة تفتت الشيء الى أجزاء كأن تكون الاجابة هى مثلا ، استعمال الصندوق فى اشعال النار ٠

ويقيس جولدنر التصلب أيضا ، فمثلا فى اختبار المكعبات يعين التصلب (ا) اذا وجد المفحوص صعوبة فى الحكم على العدد الصحيح من المكعبات٠ و (ب) اذا احتفظ بذات طريقة المواجهة بعد فشلها ٠ أو (ج) اذا يئس وتوقف دون ان يكمل حل مشكلة كما يميز التصلب بميل المفحوص الى اعطاء عدد كبير من الاستعمالات والوظائف المتشابهة منطقيا ، بينما تتميز المرونة بالتنوع فى الاستعمالات والوظائف ٠

واذا قارنا اسلوب جولدنر هذا باختبار اناء الماء الذى سبقت الاشارة اليه نلاحظ انه فى اختبار اناء الماء نحصل على درجة تمثل خاصـــة يمكن حصرها ، الجمهور مثلا ، بينما فى بطارية جولدنر تمثل كل درجة تقـديرا لعملية عقلية مفترضة ويمكن ان تظهر فى الأداء بحيث يمكن استنتاج العمليات العقلية منه ٠ وقد وجد جولدنر تاييدا كثيرا من دراساته لفرضه الذى يذهب الى عمومية الســـمتين اللتين لاحظهما أولاهما : اتجاه الكل الجـــزء ، التصلب ٠

ومن اختبارات المشكلات التى تستعمل فى قياس الشخصية اختبارات تكوين المدرك Concept-Formation فقد استعملت فى قياس الشذوذ فى العمليات العقلية ٠ ويمكن ان تأخذ اختبار هانفمان ـ كازانين Hanfmann Kasanin نموذجا لهذه الاختبارات ٠ يستعمل فى هذا الاختبار عدد

٢٢ مكعبا خشبيا ذات الوان واشكال وارتفاعات وأحجام مختلفة ، توضع على طاولة ، ومطبوع على الجانب المختص من كل مكعب مقطع عديم المعنى، وكل مقطع يميز نمطا من المكعبات ، مثلا كل المكعبات التى تحمل مقطع م ن ر تكون صغيرة وطويلة ، ويخبر المتحن المفحوص بأن المكعبات أنواع أربعة ، نوع ، منها يسمى م ن ر ، وان عليه ان يكتشف أساس التقسيم ويصـنف المكعبات ، ويعمل المفحوص على اكتشاف التصنيف كمـا يجب ، الا انه لا يسمح له بقلب المكعبات ليرى الأسماء ، ويعد كل محاولة للتصنيف يوضح المتحن خطا واحدا ثم يطلب من المفحوص أن يقوم بمحاولة أخرى ، ويستمر هذا الى أن يتم حل المشكلة ، ويعطى المبدأ الذى يقوم عليـه التصنيف · وتوضيح ملاحظة المفحوص ما اذا كان يستعمل فروضا منطقية ، مثلا ، ربما كانت كل م ن ر مستطيلات » ، أو يستعمل التخمين ، أو يستفيد من التصحيح أو القدرة على استعياب تأهب غير سليم والتخلص منه وتكوين مدرك جديد· كما يمكن ملاحظة أى أساليب أو طرق شاذة يتبعها المفحوص ·

(ب) الاختبارات الادراكية غير محددة البنية : ويمكن أن نطلق على هذا النوع من الاختبارات الادراكية ، بصفة عامة ، الاختبارات الاسقاطية أو الأساليب الاسقاطية Projective Techniqueفي قياس الشخصية · والميز الرئيسى لهذه الأساليب هى أنها تكلف المفحوص بعمل غير محدد البنية نسبيا، أي عمل يسمح بعدد غير محدد تقريبا من الاستجابات المكنة ، لهـذا فان التعليمات فى مثل هذا النوع من الاختبارات تكون عامة بحيث تترك لخيال المفحوص العنان ، وكذلك فان مثيرات الاختبار نفسه عادة ما تكون غامضة وغير محددة · والافتراض الكامن وراء هذه الأساليب أن الطريقة التى يدرك عليها المفحوص مواد الاختبار ويفسرها ، أى طريقة بنائه للموقف سـوف تعكس الجوانب الأساسية لتكوينه النفسى · أى أن مواد الاختبـار سوف تعمل فى هذه الحالة كأنها شاشة عرض يسقط عليها الشـخص أراءه واتجاهاته وأطماحه ومخاوفه وصراعاته وعدوانيته وهكذا ·

وتتميز الأدوات الاسـقاطية ايضا بانها تمثل طـرق اختبار مخفاة diseuise لأنه من النادر أن يعلم المفحوص أى نوع من التفسير النفسى سوف يعطى لاستجاباته، كما تتميز هذه الأدوات الاسقاطية بانها تتخذ اتجاها

شـــاملا global في تقدير الشخصية ، ذلك لأنها تركز على الصورة الكاملة للشخصية كلها وليس على قياس سمات منفصلة ٠

وفيما يلي نتعرض باختصار لأبرز الاختبارات الادراكية غير محددة البنيـــة ٠

١ - اختبار بنـدر - جشتالت Bender-Gestalt :

وقـــد وضـــعته لوريتـــا بنـــد في عـــام ١٩٣٨ ونشـــرته تحت اســـم (اختبـــار الجشـــتالت البصـــرى الحـــركي واســـتعمالاته الاكلينيكيـــة) ، ويشـــترك هـــذا الاختبـــار مع بعض الاختبـــارات التي سبقت الاشارة اليها في انه نبع من أبحاث مدرسة الجشتالت في الادراك ونلاحظ أن الاختبارات الادراكية بصفة عامة تقوم على اختبار كيف يتعرض الشخص للمعلومات ويعيد تنظيمها ثم يذكرها ويقررها ٠ وفي اختبار بندر تعرض أمام المفحوص أشكال ذات أنماط مختلفة من التنظيم ، نقط ودوائر صغيرة وخطوط متعرجة متقاطعة وغيرها ، ويسأل المختبر المفحوص أن ينقل مجموعة الأشكال ، ثم يلاحظ طريقته في الأداء ومدى نجاحه في نقل هذه الأشـــكال ٠

واذا كانت اختبارات الشخصية ذات البناء المحدد مصممة بحيث تقيس سمة معينة فان اختبار بندر لا يمكن أن تميزه على أساس أنه مقياس لسمة واحدة أو لمجموعة من السمات ٠ ذلك لأن استجابات الأفراد عليه قد تختلف وتتباين بدرجة كبيرة جدا ، وهذه الاســـتجابات يقوم المختبر بملاحظتها ومقارنتها وتفسيرها ٠

وقد وضعت قواعد لتصحيح اختبار بندر الا أنه عند استخدام الاختبار في أغراض تشخيصية لابد أن يعمل مصححه ومفسره على التكامل النوعي لنتائج الاختبار مع أدوات تشخيصية أخرى ٠

ومن الممكن معالجة درجات الاختبار احصائيا وذلك عن طريق ملاحظة العلامات أو الاشارات التي تميز جماعة محك معينة ، فمثلا من الممكن تحديد الاستجابات التي تميز العصابيين من العاديين والتي منها مثلا : تكديس الأشكال في نصف الصفحة فقط ، أو العد بصوت مرتفع عند نقل الأشكال ،

وهكذا · وقد اثبتت دراسات اجريت على هذا الاختبار أن هذه الاشارات وغيرها كانت مميزة فعلا بين العاديين والعصابيين (٩) ·

٢ - اختبار بقع الحبر لروورشاخ Rorchach Inkblots

وهو أكثرالاختبارات الادراكيةغيرمحددة البنيةاو الاسقاطيةأهمية واستعمالا وتعرضا للمناقشة والجدل· وقد وضعه هرمانروورشاخ Hermann Rorchach وهو سويسرى الجنسية ، فى عام ١٩٢١ · وعلى الرغم من أن بقع الحبر كانت تستعمل من قبل فى دراسة التخيل وبعض الوظائف العقلية الأخرى فقد كان روورشاخ أول من استعملها فى البحث التشخيصى للشخصية ككل ·

ويتكون الاختبار من عشرة اشكال من بقع الحبر ، خمسة منها تتضمن درجات من اللون الرمادى والأسود فقط واثنان باللون الأسود والأحمر ، اما الثلاثة الباقية فهى متعددة الألوان ، ولسهولة استخدام الاختبار وضعت الصور على بطاقات أبعادها ٧ × ٩½ بوصة ·

وعند تطبيق الاختبار تقدم للمفحوص بطاقة ثم يسال ان يتكلم عما يراه، اى عما يمكن أن تمثله بقعة الحبر ويقوم المختبر ، الى جانب تسجيل استجابات المفحوص لكل بطاقة بملاحظة زمن الاستجابة ، والوضع الذى يمسك به المفحوص البطاقة ، وملاحظات المفحوص التلقائية وتعبيراته الانفعالية ، كما يلاحظ أى سلوك عرضى يصدر عن المفحوص اثناء جلسة الاختبار · وبعد تقديم البطاقات العشر للمفحوص يسال المختبر المفحوص بطريقة منظمة عن الأجزاء والجوانب المختلفة للبقعة تلك الأجزاء والجوانب التى اعطى لها المفحوص ارتباطات · وكذلك يعطى المفحوص فرصة ليوضح ويشرح استجاباته السابقة · ويعكس الترتيب الذى تقدم به البطاقات رغبة روورشاخ نفسه فى ادخال نظام نفسى يضمن بقاء استثارة المفحوص على مستوى ممكن ·

ومنذ وضع روورشاخ اختباره · ادخلت عليه تغييرات طفيفة عن طريق باحثين اخرين ، وحدث تحول فى الاهتمام ، فبعد أن كان التركيز على التصنيف الى أنواع أمراض نفسية مختلفة ، اصبح التركيز على وصف العوامل الدينامية فى الشخصية Psychodynamics ، وقد احدث بيك

(٢٢ ، ٢٣) Beck · فى الولايات المتحدة ، تغييرات بسيطة فى تصحيح الاختبار ، ووضع له معايير أولية . الا أن كلوبفر Klopfer أحدث تغييرات أساسية فى نظام تصحيح الاختبار · وفى مصر قام سيد غنيم بدراسات استخدم فيها الرورشاخ على عينات ، كما أجرى دراسة سابقة عن مدى صلاحية الرورشاخ لقياس الذكاء (٧) ووضع مع هدى برادة معايير مصرية للرورشاخ (٧ ، ٨) · وكذلك أجرى فاروق أبو عوف دراسة مقارنة عن مدى الافادة من الرورشاخ فى التمييز بين حالات العصاب القهرى والأسوياء (١٠) ، ودراسة أخرى عن مدى صلاحية الرورشاخ لقياس الابتكارية (١١)، ودراسة ثالثة استخدم فيها المقياس فى التمييز بين القادة وغيرهم (١٢) ·

ويعتمد تصحيح الاختبار على قواعد لها طابع موضوعى ، وهناك عدد من الدرجات الرئيسية التى يمكن أن توضع ضمن الأقسام الرئيسية الثلاثة التالية التحديد المكانى ، والمحددات · والمحتوى · وتدل درجات التحديد المكانى عما اذا كانت الاستجابة تستعمل البقعة كلها (W) ، أو الى قسم فرعى يشيع ادراكه (D) ، أو تفاصيل غير مألوفة أو يندر الاشارة اليها (Dd) · اما المحددات فهى الشكل · أو اللون ، أو التظليل التى أخذها المفحوص فى اعتباره · فمثلا تعطى الاستجابة عن الحركة (M) عندما يصف المفحوص اناسا فى حركة و (CF) عندما تعتمد الاستجابة على الشكل واللون ، مع اعتبار اللون أكثر أهمية فى تحديد الاستجابة · ويلاحظ المصحح أيضا مدى ملاءمة الاستجابة لشكل بقعة الحبر ، فاذا كانت الأشكال تدرك بطريقة مألوفة تعطى (F) ، واذا كانت دقيقة بدرجة غير عادية (F +) · اما اذا كانت بعيدة الصلة ببقعة الحبر أو غير محددة ولا مقبولة فتعطى (F—) .

اما درجة المحتوى فتسجل ما اذا كانت الاستجابة تشير الى أشخاص (H) وأجزاء من أشخاص (Hd) ، أو ملابس (G) وهكذا ·

وتوجد درجة لمدى الشيوع (P) وتعطى على أساس تكرار وشيوع استجابات مختلفة بين الناس ، أما درجة الأصالة (Q) فتشير الى الاستجابات النادرة ولكن تعطى تفسيرا مقبولا لبقع الحبر ·

ويقوم تحليل استجابات الرورشاخ بعد ذلك على أساس العدد النسبى

الاستجابات الواقعة فى الأقسام المختلفة . وكذلك على النسب والعلاقات فيما بين هذه الأقسام المختلفة . ومن أمثلة التفسيرات الكيفية التى شاع استعمالها بالنسبة الى درجات الروورشاخ ربط استجابات « الكل » بمستوى التفكير ، وربط استجابات « اللون » بالانفعالية والحركة الانسانية بالتخيل وحياة الوهم . وفى الاستعمال العادى للروورشاخ يعطى تأكيد رئيسى على الوصف الكلى للشخص ، والذى ينظم الاكلينيكى فيه الناتج فى صورة متكاملة آخذا فى الاعتبار علاقات الدرجات المختلفة بعضها بالبعض الآخر ، كما يستفاد فى الواقع من اية معلومات يحصل عليها المختبر عن المفحوص ، سواء كانت من اختبارات أخرى او من المقابلات او من سجلات تاريخ الحالة ، يستفاد من هذا كله فى اعداد ذلك الوصف الكامل للشخص .

وفى الامكان وضع معايير لدرجات الروشاخ ، الا انها يجب أن تقوم على جماعات فعلية ، وعلى الرغم من أن الروورشاخ قابل للتطبيق ابتداء من سنوات ما قبل المدرسة الى سن الرشد ، فان معاييره الأساسية كانت مشتقة اساسا من جماعات من الكبار . وقد بذلت جهود حديثة لتوسع مجالات استعمال الاختبار فى مستويات اعمار متباينة (١٨ : ٥٧١) .

وقد تعرضت بعض الافتراضات التى يقوم عليها تصحيح الروورشاخ لكثير من النقد والشك نتيجة عدد متزايد من نتائج الأبحاث . فمثلا أثبتت دراسة قورنت فيها بطاقات الروورشاخ فى صورتها الملونة بالبطاقات ذاتها معروضة بحيث تلغى فيها الألوان ، أثبتت هذه الدراسة أن اللون ذاته ليس له ذلك التأثير المنسوب اليه على خصائص الاستجابات على بطاقات الروورشاخ (١٨) كما أشارت دراسات أخرى الى أن الاستعداد اللغوى يؤثر فى عدد من درجات الروورشاخ كانت تفسر عادة على انها مؤثرات على سمات معينة للشخصية ، وكذلك وجدت بعض الدراسات ارتباطا عاليا بين درجة تعقد استجابات الروورشاخ وبين درجات المفحوصين على اختبار استعداد لغوى ، وبين تعقد هذه الاستجابات والسن والمستوى التعليمى (١٨) .

ومن العوامل التى تزيد من تعقيد تفسير درجات الروورشاخ مجموع عدد الاستجابات ، والذى يعرف « بانتاجية الاستجابة » أو (R) فمن الصعوبات التى تمثلها انتاجية الاستجابة مثلا انه يبدو أن هذه الانتاجية ـ كما هو الحال بالنسبة الى تعقد الاستجابات

كما سبق ـ متصلة بعوامل مثل السن ، والمستوى العقلي ، والمستوى التعليمي ، بل ان الاكثر من هذا مغزى أن هناك ما يشير الى ان « انتاجية الاستجابة » تختلف اختلافا دالا من ممتحن الى اخر (١٨) • وتوحى هذه النتائج كلها بأن انتاجية الاستجابة والتى تعتبر فى ذاتها محددا لكثير من درجات الروورشاخ متأثرة بعوامل خارجية وبعيدة عن العوامل التى يزعم أن الروزشاخ يقيسها •

أما بالنسبة الى صدق الاختبار فان نتائج ما أجرى عليها من دراسات غير مشجعة بصفة عامة على الرغم من كثرتها ، اذ أن معظم هذه الدراسات قد فشلت فى ايجاد أية علاقة دالة بين درجات الرورشاخ وبين مجموعات من الدرجات على اختبارات أخرى أو على تقويمات عامة شاملة للشخصية• وقد وجد أن للرورشاخ صدقا تنبؤيا وصدقا مصاحبا ضعيفا عندما درس من حيث علاقته بمحكات مثل التشخيص الاكلينيكى ، وبالاستجابة للعلاج النفسى ، وبالنجاح أو الفشل فى مهن مختلفة تلعب فيها خصائص الشخصية دورا هاما ، ويوجد صراعات ومخاوف وأوهام أمكن تمييزها بصورة مستقلة • وحتى الدراسات التى أعطت نتائج ايجابية فقد اتضح أن فيها اخطاء منهجية خطيرة (٢٧ : ٥٦٤ ـ ٥٦٥) • ومما ترتب على النتائج السالبة لصدق اختبار الرورشاخ أن تغير التركيز فى التصحيح من التصحيح الادراكى التقليدى الى تحليل محتوى الاستجابات • لقد كان رورشاخ ذاته واتباعه يعتمدون على الأسس الادراكية للارتباطات التى يعطيها الشخص لبقع الحبر ، وكانت الافتراضات الكامنة وراء هذا الاتجاه هى أن استجابات المفحوص لبطاقات الرورشاخ تدل على استجاباته الادراكية العادية ، وأن سمات الشخصية تؤثر فى الادراك ، وهذا بالنسبة الى التصحيح الادراكى التقليدى ، أما فى تحليل المحتوى فان التركيز يكون على ما يدركه المفحوص فى بقع الحبر ، ومن الممكن أن ننظر الى هذا المحتوى ونعامله كما ننظر الى اى معلومات او مادة يقدمها المفحوص فى جلسة اكلينيكية • ويبدو أن النتائج فى اتجاه تحليل المحتوى واحدة وعشرة اكثر من النتائج المعتمدة على طريق التصحيح الادراكى •

فى ضوء هذه الدراسات والنتائج المتاحة حتى الآن فانه يبدو أن اختبار الرورشاخ يجب أن يستعمل على أساس انه أداة مساعدة فى المقابلة

يستعين به الاكلينكى ذو الكفـــاءة العـــالية والدربة الكافية على أعطائه وتصحيحه وتفسيره · والواقع ان الأساس الذى يقف عليه المدافعون عن استعمال الرورشاخ ـ كما يقول كرونباك ـ ـ ٢٧ : ٥٦٤) ـ هو النجـــاح الذى لقيه الاختبار فى حالات فردية ·

هذا ، ومن الطريف الذى له مغزى بعيد ان نشير بايجـــاز الى أن عوامل التفاعل الاجتماعى بين المفحوص والمختبر تؤثر فى اسـتجابات المفحوص على الرورشاخ ، فان قدرا لا بأس به من نتائج البحوث يوضح أن التفاعل الاجتماعى بين مفحوص معين ومختبر معين يؤثر على جـوانب كثيرة من استجابات المفحوص من جهة وعلى تفسير الممتحن لهذه الاستجابات من جهة أخرى (٤٥ ، ٤٦ ، ٥٧) ، وهذا مجال خصب وفسيح للدراســـات العربية تكمل وتنمى ما بدأه الباحثون فى الغرب من عمل جـاد حـــول الرورشاخ ·

ونحب أن نختتم هذا العرض الموجز لاختبار الرورشاخ بنظرة الى قدرة الاختبار على تمييز السواء أو جوانب القوة الشخصية ، ذلك لأن هناك قدرا كبيرا من نتائج البحوث يشير الى ان أوصاف الشخصية المعتمدة على سجلات الرورشاخ منحازة الى اتجاه عدم التوافق أو عدم السواء · ويرجع والين Wallen (٦٠ : ٢١٤) هذا الى أن طريقة الرورشاخ نشأت وتطورت فى معظمها متصلة بالأشخاص غير الأسوياء ، والاحتمالات الراجحة ، فى رأى والين ، أن هذه الطريقة ليست مناسبة لتحديد جوانب القوة والسواء فى الشخصية · ويستدل والين على هذا بالدراسة المستفيضة التى قامت بها رو Roe (٥٣) والتى استعملت فيها الرورشاخ لدراسة بعض العلماء البارزين ، والتى قررت فيها عجبها من أن سجلات الرورشاخ التى جمعتها عن هؤلاء العلماء لم تعطها أية اشارة الى الحيـــاة الواقعية أو الأداء الواقعى لهؤلاء العلماء المفحوصين · وهى تشير الى أن هذه الســـجلات توضح فى الحقيقة وجود مشكلات جادة فى حيـــاة هؤلاء المفحوصـــين ولكنها ، أى السجلات لا تعطى ما يوضح أن هؤلاء المفحوصين استطاعوا ان يواجهـــوا هذه المشكلات بنجاح وفاعلية · وان اتجاه استعمال الرورشاخ فى دراســة هذه الجوانب السوية والقوية فى الشخصية يمثل أيضا مجالا جديرا باهتمام الباحثين العرب ·

٣ ـ اختبار بقع الحبر لهولتزمان

The Holtzmann Inkblot Technique

صمم هذا الاختبار على نمط الرورشاخ ، الا ان هولتزمان (٣٧) حاول فيه أن يتجنب العيوب الفنية الرئيسية الرورشاخ . والتغيرات التى ادخلها هولتزمان على الاختبار سواء فى مواد المثيرات أو الطريقة تغيرات كبيرة وشاملة تجعل من الممكن أن ندرسه ونقومه دون الاشارة الى الخصائص النظرية أو العملية للرورشاخ . ويتضمن اختبار هولتزمان مجموعتين متماثلتين من سلاسل البطاقات ، تضم كل مجموعة عدد ٤٥ بطاقة . والمطلوب من المفحوص أن يعطى استجابة واحدة لكل بطاقة وعلى هذا نثبت انتاجية الاستجابة بالنسبة الى كل مفحوص ونتلافى كثيرا من عيوب التصحيح التقليدى للروشاج وطريقة اجراء الاختبار وتصحيحه مقننة تقنينا جيدا ومشروحة بشكل واضح . وللاختبار معايير لعينات سوية تبدا اعمارها من الخامسة الى الرشد ، وعينات غير سوية ، ويبدو ثبات المصحح مقبولا ، ونتائج الثبات ، سواء الاتساق الداخلى ، أو التكافؤ مشجعة ، وان كانت دراسات الصدق ما زالت قليلة . وللاختبار صورة جماعية تعطى درجات قريبة من الدرجات التى يعطيها الاختبار عندما يطبق بصورة فردية .

٤ ـ اختبار تفهم الموضوع .

Thematic Apperception Test (TAT)

وهناك اختبار آخر من الاختبارات الادراكية غير محددة البنية ، أو الإسقاطية ، الا أنه يختلف عن الاختبارات التى سبقت مناقشتها فى أن مثيراته تعتبر اكثر تحديدا فى البنية نسبيا ، ولأنه يتطلب استجابات وانشطة اكثر تعقيدا واكثر خضوعا للتحكم العقلى من مثيلاته من الاختبارات السابقة . ويعتمد تفسير هذه الاستجابات عادة على تحليل محتوى ذى طابع كيفى وذلك تمشيا مع الاتجاه الغالب الى أنه من الأكثر فائدة أن نفسر الاختبارات الاسقاطية عامة على أساس تحليل المحتوى ، ويعطى اختبار تفهم الموضوع وغيره من الاختبارات التى تعتمد على القصص مجالا أكثر اتساعا لتحليل المحتوى من اختبار للرورشاخ .

وكان أول ما نشر عن اختبار تفهم الموضوع مقال نشره هنرى مورى H. Murray وزميلته Morgan فى عام ١٩٣٥ ، واستمرت الأبحاث التى قام بها مورى وزملاؤه فى العيادة النفسية فى جامعة هارفارد بالولايات المتحدة ثم ظهرت فى كتاب فى عام ١٩٣٨ (٥٠) وقد استعمل هذا الاختبار منذ ظهوره استعمالا واسعا فى الممارسة الاكلينيكية وفى البحث ، كما يعتبر نموذجا اقتدى به كثير من اختبارات الشخصية ٠

تتكون مواد الاختبار من ١٩ بطاقة تحتوى كل منها صورة غامضة باللونين الأبيض والأسود ، بالاضافة الى بطاقة خالية من أى صورة ٠ ويطلب من المفحوص أن يكون قصة تناسب كل صورة ، ذاكرا ما أدى الى ما تمثله الصورة ، واصفا ما يحدث وما تحس به وتفكر فيه الشخصيات الممثلة فى الصورة ، وعلى أن يذكر النتيجة التى يتوقع أن ينتهى اليها ما يصفه ٠ أما بالنسبة الى البطاقة البيضاء التى ليس عليها أى صورة ، فان التعليمات تطالبه بأن يتخيل صورة ما على البطاقة ويصفها ثم يذكر قصة عنها ٠ وتتطلب الطريقة الأصلية كما حددها مورى جلستين لكل منهما ساعة ، مع استعمال ١٠ بطاقات فى كل جلسة ، وتختار البطاقات المحجوزة للجلسة الثانية على أساس أنها غير مألوفة وشاذة وذات طابع درامى ، والتعليمات المصاحبة هى أن يطلق المفحوص لخياله العنان وتوجد أربعة مجموعات من الاختبار ، وهى متداخلة ، وتتكون كل مجموعة من ٢٠ بطاقة ، منها مجموعة للاولاد ، وأخرى للبنات ، وثالثة للذكور فوق سن ١٤ ، ورابعة للاناث فوق سن ١٤ ، ويستعمل معظم الاكلينيكيين مجموعات مختصرة من بطاقات منتقاة ، ونادرا ما يعطى أكثر من عشر بطاقات لمفحوص واحد ٠

وعلى الرغم من أن الاختبار أصلا هو اختبار فردى شفوى يعطى فى الموقف الأكلينيكى ، فمن الممكن اعطاؤه كتابة ، أى يقوم المفحوص فيه بالكتابة بدل الكلام كما يمكن اعطاؤه فى جماعة ٠ وتشير نتائج بعض البحوث الى أنه فى حالة اعطائه جماعيا وكتابة يكون انتاج المواد ذات المعنى والمغزى أكثر سهولة ويسرا (١٨ : ٥٠١) وهناك ميل الى تأكيد أهمية العوامل والمؤثرات الاجتماعية فى الموقف الاختبارى بالنسبة الى اختبار تفهم الموضوع ٠ كما لاحظنا أثر هذه العوامل والمؤثرات ونوهنا بها بالنسبة الى اختبار

الرورشــاخ ، ونكـــرر ان هذا أيضا مجـــال طيب لمزيد من البحث النفسى الاجتمــاعى ٠

اما فى تفسير قصص اختبار تفهم الموضوع فيحدد المختبر أولا البطل hero وهو عبارة عن الشخصية ، ذكرا كانت ام أنثى ، التى يتوحد معها المفحوص ثم يحلل محتوى القصص اساسا وفق ما حدده من حاجات needs وضغط Press ٠ وقد سبقت الاشارة الى الحاجات عند مورى عند الكلام عن مقياس التفضيل الشخصى الذى وضعه ادواردز ، ومن امثلة هذه الحاجات الانجاز achievement والاعتداء aggression ، والاعالة nurturance ، والجنس ٠ الا أن مورى يعطى حاجات معينة أهمية كبيرة فى التفسير مثل : تجنب الأذى، البحث عن الكمال، الاعتداء، السيطرة، العدوان الموجه نحو الذات ، الاستنجاد ، السلبية ، الجنس (٩ : ١٤٤ وما بعـــدها) ٠

ومع معرفة البطل الرئيس ، ودراسة الحاجات والدوافع الأساسية للسلوك لابد من معرفة ما اســماه مورى الضــــغط أو الضغوط ، وهى تشير الى القوى البيئية التى قد تيسر أو تعوق اشباع الحاجات ، ومن أمثلة هذه الضغوط ما يشار اليه من التعرض للهجوم أو النقد بواسطة شـخص أخر ، وتقبل العطف والمحبة ، والتعرض للخطر المادى أو الجروح المادية ٠

وعند تقدير أهمية أو قوة حاجة أو ضغط معين عند الشخص يعطى اهتمام خاص بعمق وحدة وتكرار حدوث هذه الحاجة أو الضغط فى القصص المختلفة بالاضافة الى الاهتمام بارتباط الحاجة أو الضغط بصورة معينة بصفة خاصة ٠ والافتراض القائم هو أن أية استجابات مخالفة للاستجابات الشائعة لكل صورة يغلب أن يكون لها _ أى الاستجابات المخالفة _ مغزى بالنســـبة الى المفحوص ٠

وعلى الرغم من أن تفسير اختبار تفهم الموضوع كان أساسا تفسيرا كيفيا وانطباعيا كما سبقت الاشارة ، فان كرونباك (٢٧ : ٥٧١) يذهب الى أنه من الممكن انشاء وتطوير نظم موضوعية لتفسير هذا الاختبار ، وفى رأيه أن هناك عددا كبيراً من المتغيرات العامة المشتركة التى يمكن أن نلاحظ قوتها

تقريبا فى كل اداء على الاختبار ، ومن امثلتها : ادراك السلطة ، ورد الفعل للاعمال الصعبة جدا ، والاصالة : والاعتماد على الحظ وتأثير السحر ·
وتوجد امثلة متعددة لمحاولات التصحيح الكمية التى اعطت ثباتا عاليا للمصححين ، الا أن تطبيق هذه الأساليب يستغرق وقتا طويلا لهـذا ندر استعمالها فى المجال الاكلينيكى ·

وقد نشر قدر لا بأس به من المعلومات المعيارية عن الاختبار فيما يتصل باكثر الاستجابات المميزة لكل بطاقة شيوعا ، وبالطريقة التى تدرك بها كل بطاقة ، والموضوعات المذكورة ، والأدوات التى تنسب الى الشخصيات ، وسرعة الاستجابات ، وطول القصص ، وعلى الرغم من أن هذه البيانات الخاصة بالمعايير تمثل اطارا عاما لتفسير الاستجابات المفردة ، فان معظم الاكلينيكيين يعتمدون اعتمادا رئيسيا على « معايير ذاتية » بنوها من خبرتهم الشخصية فى اعطاء الاختبار ·

وقد استعمل اختبار تفهم الموضوع اسـتعمالا واسعا فى بحـوث الشخصية ، ومن أمثلة استعماله فى مصر بحث احمد عبد العزيز سلامة (٣)· وركز بعض هذه البحوث على الافتراضات التى تقوم عليها تفسيرات الاختبار مثل توحد الذات مع البطل ، والمغزى الذاتى للاستجابات غير المألوفة · وعلى الرغم من أن هذه البحوث لا تعطى تأييدا للصدق التنبؤى والصـدق التلازمى ، فانها تضيف الى صدق التكوين الفرضى لتفسيرات اختبار تفهم الموضوع · ومن الافتراضات الأساسية الأخرى التى يشارك فيها هذا الاختبار الاختبارات الاسقاطية الأخرى أن ظروف المفحوص العاطفية والدافعية وقت الاختبار تؤثر على استجابات المفحوص لموقف الاختبار غير محدد البنية · ويتوافر قدر كبير من نتائج الدراسات التجريبية يشير الى أن ظروفا مثـل الجوع ، والحرمان من النوم · والحبوط الاجتماعى · ومعاناة الفشل فى موقف اختبارى سابق ، كلها تؤثر على استجابات اختبار تفهم الموضوع · وبينما يؤكد هذا صحة الفرض الاسقاطى فان حساسية هذا الاختبار لهـذه العوامل المؤقتة يزيد من صعوبة الكشف عن سمات الشخصية الثـابتة (١٨ : ٥٠١) ·

ولعلنا نلاحظ أن اختبار تفهـم الموضـوع قد صمم اسـاسا بحيث

يشمل جوانب الشخصية جميعا ، كل الأفكار والمشاعر والسلوك ، وعلى هذا فلا يستطيع ان يسبر موضوعا أو جانبا واحدا بعمق ، ولهذا فقد انشئت اختبارات تفهم موضوع تركز على جانب بعينه ، اى انها تصمم لاستدعاء استجابات تدور كلها حول مسألة واحدة وقد بداها مورى وليكرت Murray and Likert فى دراسة بين العمال والادارة وذلك عن طريق تقديم صورة تمثل تظاهرات مع رجال الأمن ٠

اما المحاولة المستفيضة المتعمقة للاستفادة من هذه الاختبارات المركزة فى تفهم الموضوع فكانت محاولة ماكليلاند ورفاقه (٤٧ ، ٤٨) وكذلك محاولة اتكنسن (١٩ ، ٢٠) ٠

لقد اختار ماكليلاند اربع صور ، منها اثنتان من اختبار تفهم الموضوع، كان القصد استدعاء الاتجاهات نحو الانجاز ٠ وكانت هذه البطاقات الأربع تمثل موقف عمل (رجل امام ماكينة) ، وموقف مذاكرة (تلميذ يجلس الى مكتبه ومعه كتاب) ، وصورة أب وابن ، وصورة ولد يبدو مستغرقا فى احلام اليقظة ٠ ويقاس دافع الانجاز فى كل قصة تعبر عن التنافس والكفاح للوصول الى مستوى أو تحقيق هدف اما اتكنسن فقد استعمل اختبار اخر هو French Test of Insights ولذات الغرض ، أى قياس الدافع الى الانجاز ، ويقوم الاختبار على اعطاء وصف موجز للسلوك مثل : « يعطى الفتى بل فرصته لزميله ليفرز » ، ويسأل المفحوص ان يعطى تفسيرا لهذا ٠ ويتضمن الاختبار عددا من مثل هذه العبارة وتتكون الدرجة عليه من عدد المرات التى تذكر فيها الرغبة فى الانجاز باعتبارها دافعا ٠ وتمثل هذه الدراسات خطا من الأبحاث الغنية التى نبعت أصلا من اختبار تفهم الموضوع ٠

٥ ـ اختبار تفهم الموضوع للاطفال :
Children's Apperception Test (CAT)

وعلى الرغم من أن اختبار تفهم الموضوع الأصلى وضع بحيث يصلح فى التطبيق على أطفال يبلغون من العمر اربع سنوات فان اختبار تفهم الموضوع

للاطفال قد صمم خاصة ليطبق على الأطفال ممن تتراوح اعمارهم بين ثلاث وعشر سنوات • ويستعمل هذا الاختبار على بطاقاته صور الحيـوانات بدل البشر على أساس افتراض أن الأطفال يسقطون على الحيوانات بسهولة اكثر من اسقاطهم على البشر ، الا أن صور الحيوانات موضوعة في مواقف بشرية، وهذه الصور مصححة بحيث تستدعى عند الأطفال أوهاما وتخيلات تتصل بمشـكلات التغــذية والأنشــطة الفنيـة الأخـرى ، ومشــكلات العلاقات بين الأشقاء ، والعلاقات مع الأبوين ، والعدوان ، وعمليــات الاخـراج وضبطها ، والخبرات الأحرى الخاصة بالطفولة ، وعلى العكس من افتراض واضعى هذا الاختبار ، اثبتت دراسات متعددة أن استعمال الصور البشرية وصور الحيوانات لم يعط اختلافات تذكر ، بل ظهر غالبا أنه فى حالة استعمال الصور البشرية كانت انتاجية الأطفال من المواد ذات المغزى الاكلنيكى اكبـر منها فى حالة استعمال صـور الحيوانات ، وفى ضوء هذه النتائج صـمم واضعو الاختبار صورة بشرية منـه ، هى CAT-H ، للاستعمــال مع الأطفال الكبار ، وخاصة من يزيد عمرهم العقلى عن عشر ســنوات ، وذهبوا الى أن أيا من الصورتين ، البشرية والحيوانية ـ يكون اكثر فاعلية حسب سن الطفل وخصائص شخصيته •

(ج) الاختبارات اللفظية غير محددة البنية :

على الرغم من الاختبارات غير محددة البنية التى نوقشت حتى الآن تتطلب استجابات لفظية بينما المثيرات غير لفظية ، فان هنــاك نوعا من الاختبارات غير محـددة البنية أو الاســقاطية تعتمد على مثيرات لفظيــة كما تتطلب اســتجابات لفظية ســواء كانت منطوقة أو مكتــوبة ، وهى مناســبة للتطبيق الجماعى بالكتابة • وتذكر فيمــا يلى مثــالين من هــذه الاختبارات •

١ ـ اختبار تكملة الجمـل :

وتستعمل اختبارات تكملة الجمل فى اختبارات الاستعداد اللغــوى الا انها اختبارات تكملة الجمل ذات البنية غير المحددة تختلف عن الأولى فى انها تتطلب استجابات وتكملات متنوعة ومتباينة على أساس انها نـوع من

الاختبارات الاسقاطية . وغالبا ما تعطى الكلمة الأولى ، أو بدايات الجمل ، وعلى المفحوص أن يكتب النهاية ، ومن امثلتها ما يلى :

انا احس ...
ما يضــايقنى ...
النســاء ...

وتصاغ بدايات الجمل بحيث تستدعى اســتجابات تتصل بمجــال الشخصية المراد التعرف عليه ، وتعتبر هذه المرونة التى يتمتع بها اسلوب تكملة الجمل بميزة هامة تزكيه للاستعمال الاكلينيكى وفى البحث . وقد نشرت صور مقننة من هذه الطريقة للاستعمال فى مجال اوسع ، منها قائمة روتر للجمل الناقصة Rotter Incomplete Sentense Blank

وتتكون من عدد ٤٠ بداية للجمل ، وتنص التعليمــــات الموجهة الى المفحوص على أن يكمل هذه الجمل ليعبر عن مشــاعره الحقيقية ، والا يترك اية جملة ويتأكد من انه يكون جملة كاملة ، ومن أمثلة هذه الجمل الناقصة .

انا احب ..
ان خوفى الأكبر ..
احيـانا ...
انا ارغب ...
والدى ...

وتقدر كل تكملة على مقياس ذى سبع نقط حسب درجة التوافق أو عدم التوافق التى تدل عليها التكملة . ويعطى دليل الاختبار عينة من التكملات والدرجة المناسبة لها معا يسمح بقدر مناسب من التقدير الموضوعى للدرجة، ويمثل مجموع التقديرات درجة كلية للتوافق يمكن استعمالها فى اغــراض الانتقاء . ومن الممكن ان يفحص محتوى الاستجابة للتعرف على ما قد يكون فيها من مؤشرات ذات دلالة تشخيصية ، وتشير دراسات صدق الاختبار الى امكانياته الواسعة فى الاستعمال .

٢ ـ اختبار تداعى الكلمات : استعمل تداعى الكلمات فى مراحل علم

النفس التجريبى والقياس العقلي المبكرة للتعرف على عمليات التفكير ، ثم اتسع مجاله فى الميدان الاكلينيكى عندما انتشر أسلوب التحليل النفسى ولعل من اهم الاختبارات المعتمدة على تداعى الكلمات والتى تزايدت اهميتها منذ وضع اختيار كنت ـ روزانوف للتداعي الحر Kent — Rosanoff Free Association Test والـــذى وضـــع فى عـام ١٩١٠ ، واعتمــد تمـاما على التصحيح الموضـــوعى وعلى المعـــايير الاحصائية · وتتكـون الكلمـــات المثيـــرة من مائة كلمـــة محـــايدة شـــائعة اختيـــرت لانها تميـــل الى اس‍ـ‍‍ـتدعاء مثيـرات متشـابهة عنـــد النـاس بصفة عامة · مثلا يميل معظم النـــاس الى أن يستجيبوا لكلمـة طاولة بكلمة كرسى ، ولكلمة ظلام بكلمة نور ، واعدت مجموعة من جداول التكرار ـ واحد لكل كلمة مثير ـ مبينة عدد المرات التى أعطيت فيها الاستجابة لعينة تقنين عددها الف راشد سوى ·

وفى تصحيح اختبار كنت ـ روزانوف ، تستعمل قيمة وسيط التكرار للاستجابات التى أعطاها المفحوص للمائة كلمـــة المثيرة على أنها « مؤشر للشيوع » ، وتعتبر الاستجابات التى لا توجد فى جداول المعايير « فردا » ، وقد بينت مقارنات الذهانيين بالأسوياء أن الذهانيين أعطوا عدد استجابات « فرد » أكثر ، كما حصلوا على « مؤشر للشيوع » منخفض عن الأسوياء · الا أن القيمة التشخيصية للاختبار بدأت تتناقص مع التحقق التدريجى من أن تكرار الاستجابات يختلف باختلاف السن، والمستوى الاجتماعى الاقتصادى والمستوى التعليمى وغيرها من العوامل · وعلى هذا فان التفسير السليم للدرجات على الاختبار يتطلب معايير الجماعات الفرعية ، وكذلك معلومات اضافية عن المفحوص · وعلى اية حال فان الاهتمام بهذا الاختبار مستمر ومتجدد فى مجالات السلوك اللغوى والشـــخصية فضلا عن اســـتعمالاته الاكلينيكية · وقد قام فؤاد ابو حطب بنقل هذا الاختبار الى اللغة العـــربية وطبقه على عينات مصرية بلغ حجمها ١٠٠٠ مفحوص واعد له معايير تقبل المقارنة مع المعايير الأجنبية (*) ·

(*) فؤاد ابو حطب : المعايير المصرية لاخبار تداعى الكلمات : فى بحوث تقنين الاختبارات النفسية ، المجلد الاول (تحرير فؤاد ابو حطب) ، القاهرة ، مكتبة الانجلو المصرية ، ١٩٧٧ ·

الفصـــل السادس عشر

قيـاس الميـول

يرجع الاهتمام بالدراسة النفسية للميول الى التوجيه التربوى والمهنى، كما يرتبط انشاء الاختبارات التى تقيس هذه الميول بعمليتى الانتقاء والتصنيف المهنى ، لأن للميول اهمية عملية ، ويجب أن تؤخذ فى الاعتبار من جانب كل من طالب العمل وصاحب العمل ، فالأداء فى عمل من الأعمال هو خلاصة للاستعداد وللميل معا • وعلى هذا فان قياس هذين المتغيرين ، أى الاستعداد والميل ، يسمح بدرجة اكبر من الدقة فى التنبؤ بالأداء من الاقتصار على واحد منهما فقط • ذلك لانهما ، أى الاستعداد والميـــل ، على الرغم من ارتباطهما فان مستوى مرتفعا فى أحدهما ليس يعنى بالضرورة مســـتوى مرتفعا فى الآخر • فقد يكون عند الشخص استعداد للنجاح فى نمط معين من النشاط ، سواء كان دراسيا ، أو مهنيا ، أو ترفيهيا ، دون أن يكـون عنده الميل المناسب لهذا النشاط أو قد يكون عنده ميل لعمل ما ولكنه يفتقـد الاستعدادات المطلوبة لهذا العمل (راجع ١ : ٢٧٥ وما بعدها) • اما وقد تكلمنا من قبل عن قياس الاستعدادات ، فاننا نتناول فى هذا الفصل قياس هذا المتغير الآخر وهو الميل •

مقاييس الميـول

يعتبر اختبار الميول استفتاء أو استبيانا مطولا يستخدم اسلوب التقرير الذاتى الذى تكلمنا عنه فى الفصل الثامن ، ويهدف هـــذا الاستبيان الى الحصول على معلومات بأن نجعل الشخص يصف خصائصه ومميزاته الذاتية وهذا الاختبار أو الاستبيان عبارة عن نوع من المقابلة المكتوبة ، ولما كـان يتضمن عددا كبيرا من الأسئلة غير المباشرة • فانه يفضل المقابلة الشـفوية المباشرة •

والواقـع ان طـريقة السؤال المبـــاشر للمفحوص لتتعرف على ميله• تعنى ان تساله عما يميل اليه من اعمال أو دراسات او ترويح ، وهذه الطريقة

تعرضت لنقد شديد في وقت مبكر ، في خلال العقد الثاني من القرن العشرين • فقد تجمعت دراسات شتى بينت أن الاجابة عن الأسئلة المباشرة فيما يتصل بالميول كثيرا ما تكون غير ثابتة ، وسطحية ، وغير واقعية (١٨) • وترجع أسباب هذا الى أن معظم الناس ليس لديهم المعلومات الكافية عن الأعمال ومواد الدراسة والأنشطة المختلفة المتضمنة في عمل ما ، ومن ثم فانهم يكونون غير قادرين على أن يقرروا ما اذا كانوا سوف يحبون كل ما يشتمل عليه ما يحبونه من نشاط • وقد يكون ميلهم ، أو عدم ميلهم ، بالنسبة الى عمل من الأعمال راجعا الى قصور ما لديهم من معلومات عما يتضمنه هذا العمل من تفاصيل واجراءات وتصرفات تجرى كل يوم • ومن أسباب عدم الثبات والسطحية وعدم الواقعية هذه شيوع صور وأفكار جامدة ثابتة • عن بعض الأعمال والمهن ، كما هو الحال بالنسبة الى المدرس والمهندس والمحامي والطبيب ، وغيرهم ، هذه الصور والأفكار تشيع وخاصة بين الشباب ، عن طريق وسائل الاعلام المختلفة • فالمشكلة اذن هي أن الفرد ليس في وضع يسمح له بأن يعرف ميوله في المجالات المختلفة قبل أن ينخرط في سلك العاملين في هذه المجالات ، وقبل أن تفوت الفرصة للاستفادة من معرفته حين لايستطيع تغيير عمله أو تبديله • لهذا كله عدل عن طريقة السؤال المباشر وتحول الاتجاه الى الاعتماد على الأسئلة غير المباشرة المتضمنة في اختبارات الميول التى سبق وصفها بانها بمثابة المقابلة المكتوبة •

وقد اتبعت في بناء اختبارات الميول وتصحيحها عدة اساليب :

١ ـ فمنها ما أعتمد في بناء تكوينه على اساس تجـريبي emperical واقعي ، مثل اختبار سترونج للميول المهنية •
Strong Vocational Interest Blank (SVIB)

٢ ـ ومنها ما اعتمد على التنوع والتعدد ، مثل اختبار كيودر •
Kuder Preference Record

٣ ـ ومنها اخيرا ما اعتمد على اساس منطقي مثـل اختبـار لي ـ ثورب •
The Lee — Thorpe Inventory

وسوف نتناول كلا من هذه الاختبارات •

اولا : اختبار سترونج للميول المهنية :

نشا اختبار سترونج مع مجموعة أخرى من اختبارات الميول التى انبثقت عن سيمنار انعقد فى العام الدراسى ١٩١٩ ـ ١٩٢٠ فى مؤسسة كارينجى للتكنولوجيا فى الولايات المتحدة Carnegie Institute of Technology . حيث كان موضوع هذا السيمنار هو الميول ، الا انه لم يكتب لأي من تلك الاختبارات الاستمرارية والازدهار مثل ما كتب لسترونج الذى تعرض منذ نشاته للمراجعة والاتساع والبحث · وقد قام بترجمته الى اللغة العربية عطية محمود هنا ·

وتتكون الصورة الحالية لاختبار سترونج ، فى اصله الأجنبى ، والمنشورة فى عام ١٩٦٦ من عدد ٣٩٩ عنصرا مجمعة فى ثمانية اجزاء · وفى الأجزاء الخمسة الأولى يسجل المفحوص تفضيله بان يضع دائرة حول احد الحروف L (يحب Like) أو I (محايد Indiffirent) أو D (لا يحب Dislike) بالترتيب · وكل واحد من هذه الأجزاء الخمسة مختص بواحد من الأقسام التالية : المهن ، المواد الدراسية ، الأنشطة المسلية، وانشطة متنوعة (مثل اصلاح الساعات · او التعبير عن الآراء والأحكام علنا بصرف النظر عن نقد الآخرين ···) والخواص الغريبة وغير المألوفة للناس · اما الأجزاء الثلاثة الأخرى من الاختبار فتتطلب من المفحوص ان يرتب أنشطة معينة حسب تفضيله ، ويقارن ميله فى ازواج من العناصر ، وأن يقدر قدراته الحالية وخصائصه ومميزاته ·

يصحح الاختبار بمفتاح خاص لكل مهنة ، وحاليا يوجد ٥٤ مقياسا مهنيا لتصحيح الصورة الخاصة بالرجال من الاختبار و ٣٢ مقياسا مهنيا لتصحيح الصورة الخاصة بالنساء ، وتنشا مقاييس جديدة من وقت لآخر كلما تجمعت معلومات وبيانات عن جماعات عمل او مهن أخرى وعند انشاء هذه المقاييس المهنية قورنت استجابات أشخاص ناجحين فى مهنة باستجابات الرجال بصفة عامة أو النساء بصفة عامة · وتتكون جماعة المرجع هذه على أساس افتراض هو ان معظم مقاييس

المهن فى اختبار سترونج تتصل بالمهن والوظائف التجارية العليا • الا أنه قد ثبت أن استخدام جماعة مرجع ممثلة بكل مجتمع الذكور كان اقل نجاحا فى التمييز بين المهن المفردة • فميول المهنيين ورجال الأعمال كجماعة تختلف اختلافا شاسعا عن ميول العمال المهرة ، فعندما استعملت جماعة المرجع العامة ادى هذا الى طمس وعدم وضوح الفروق بين المهن ذات المستوى المرتفع •

وقد اختبرت عناصر الاختبار لكل مقياس مهنى وأعطيت لها أوزان على اساس اختلاف تكرار اختيار الرجال فى المهنة المتصلة بالمقياس والرجال بصفة عامة • فمثلا فى مقياس المحامى يدل وزن + ١ على أن الاستجابة تعطى اكثر فيما بين المحامين عنها بين الرجال عموما ، بينما يدل وزن ـ ١ على أن الاستجابة تعطى اقل فيما بين المحامين عنها بين الرجال عموما ، والاستجابات التى لا تميز بين المحامين والرجال بصفة عامة ، لا تظهر فى مقياس المحامى • بصرف النظر عن مقدار اختيار المحامين لها • وتتكون الدرجة الخام للمفحوص على أن كل مقياس مهنى من المجموع الجبرى لأوزانه الموجبة والسالبة وتحول هذه الدرجات الخام الى درجات معيارية متوسطها ٥٠ وانحرافها المعيارى ١٠ حسب توزيع الدرجات فى كل جماعة محك مهنية occupational criterion group وجماعات المحك المهنية هذه كانت تتكون عادة من ٣٠٠ شخصا من الرجال تتراوح اعمارهم بين ٢٥ و ٥٥ ، ملتحقين باعمالهم لمدة ثلاث سنوات أو اكثر ، وقد قرروا أنهم راضون عن اعمالهم •

وقد أعطيت تقديرات حروف لزيادة سهولة وتفسير الدرجات المعيارية، فمثلا التقدير (أ) (A) يمثل درجة عند او اكبر من ـ ½ ع (انحراف معيارى)، اى نقطة فوقها نحو ٦٩٪ من جماعة المهنة المعينة قد أعطوا استجاباتهم • وبالمثل فان الحد الأدنى لتقديرات ب (B) قد وضع عند ـ ٢ ع ، وتقدر الدرجات أقل من ـ ٢ ع بالتقدير حـ (C) • وعلى هذا فان التقدير حـ يعنى ان الشخص قد حصل على درجة على هذا المقياس بالذات أقل من ٩٨٪ من العينة المماثلة • وعلى اى حال فانه جدير بالملاحظة أنه مهما كانت الصورة التى تمثل بها الدرجات على اختبار سترونج ، فان هذه الدرجات معتمدة على الاختلافات بين ميول جماعة مهنية وميول الرجال عموما • وعلى هذا فان درجة مرتفعة على مقياس الطبيب مثلا تدل على أن ميول الشخص تختلف

عن ميول الرجـال بصفة عامة فى نفس الاتجـاه الذى تختلف فيه ميـول الأطبـاء ٠

هذا هو المقصود من قولنا ان سترونج قد بنى اختباره على أسـاس تجريبى عملى صرف (٢٧ : ٤٠٦) ، دون أن يعتمد الا على قدر ضئيل جدا من المسلمات والافتراضات النفسية ، وأقام تصحيحه للاختبار ـ كما لاحظنا ـ على أساس الارتباطات الملاحظة للاستجابات مع المحك ٠

ومن الممكن أن يصحح اختبار سترونج بالنسبة الى الفرد لمهنة واحدة: فمثلا عندما تهدف الى الانتقاء فاننا نحب أن نعرف الى أى مدى تتشابه ميول طالب الوظيفة مع أولئك الناجحين من العاملين فيها ، الا أن الاختبار غالبا ما يصحح كله مما يعطى فرصة لدراسة النمط الكلى لميول المفحوص ، ويعطى هذا النوع من التحليل فرصة أوثق للتنبؤ بالميل المهنى الغالب ٠

ويتضمن اختبار سترونج ، الى جانب المقاييس المهنية ، أربعة مقاييس غير مهنية ، هى :

مستوى التخصص	Specialization Level (SL)
ومستوى المهنية	Occupational Level (OL)
والذكورة والأنوثة	Masculanity — Feminity (MF)
والتحصيل الأكاديمى	Academic Achievement (AACH)

وقد انشىء مقياس مستوى التخصص بمقارنة استجابات المتخصصين فى الطب باستجابات جماعة من الأطباء عامة ٠ وقد وجد أن هذا المقياس يمكن تطبيقة فى مجالات أخرى غير الطب فى التعرف على أولئك الذين سوف يجدون متعة فى الدراسة المتقدمة التى تتضمن تخصصا دقيقا ٠ أما مقياس مستوى المهنة فيقيس الاختلاف بين ميول العمال غير المهرة فى ناحية ورجال الأعمال والمهنيين فى الناحية الأخرى ٠ أما مقياس الذكورة والأنوثة فانه يوضح درجة التشابه بين ميول الشخص وميول الرجال والنساء ٠

اما مقياس التحصيل الأكاديمى ٠ وهو أحدث اضافة الى المقاييس غير

المهنية ، فقد انشىء بمقارنة استجابات جماعات من طلاب المدارس الثـانوية الجامعة ممن حصلو اعلى درجات مرتفعة بجماعات ممن حصلوا على درجات منخفضة · وقد تبين من البحوث التى تلت اضافة هذا المقياس أنه يعكس ميولا للاجتهادات العلمية والعقلية أكثر من ميول للبيع أو الأعمال التجارية (١٨ : ٤٧٠) ·

وقد تعرض اختبار سترونج لبرنامج مستمر من البحث الذى أعطى قدرا كبيرا من البيانات المتصلة بثباته وصدقه فقد كانت معاملات الثبات ـ بالتجزئة النصفية على أساس أرقام العناصر الزوجية والفردية ـ لكل مقياس فرد ٨٨ر٠ ، وكان واحد منها فقط أقل من ٨٠ر٠ وكذلك كانت معـــاملات الاستقرار باعادة الاختبار بعد فترة أسبوعين وكذلك بعد فترة ثلاثين يوما عالية ايضا ، كما أثبتت دراسات طولية بدرجة استقرار مناسبة لفترة طويلة للدرجات على المقاييس المفردة · وكانت هذه الدراسات تمتد بين ثلاث سنوات وثمان سنوات ، وامتد بعضها الى ٣٠ سنة أو أكثر (٥٨) ·

اما صدق اختبار سترونج ، فيوجد قدر كاف من نتائج البحوث يبين أن المفحوصين يميلون الى الالتحاق والاستمرار فى المهن التى حصلوا فيها على درجات مرتفعة على الاختبار ·

وبالنسبة الى تلاميذ الثانوى فانه من الممكن التنبؤ بدرجة حسنة بالمجال أو جماعة المهنة التى سيلتحق بها التلميذ · أما التنبؤ بمهنة بعينها فانه لا يوثق به تماما الا فى حالة تلك القلة من التلاميذ التى تظهر عندها أنمـــاط محددة ومميزة جدا من الميـول ·

ويتوافر قدر من البيانات التى توضح وجود ارتباط بين درجات اختبار سترونج والنجاح فى المهنة أو الرضا عنها ، اما الأدلة التى تثبت عكس هذا فهى نادرة ويصعب تفسيرها ·

والواقع أن المزايا التى يتمتع بها اختبار سترونج جعلت انستازى تذهب فى تقديره الى القول بأنه « واحد من انجح الطرق لقياس المتغيرات

غير العقلية ، (١٨ : ٣٢٦) • وما زال الاختبار خاضعا لعدد من مشروعات البحث فى « مركز البحث فى قياس الميول » •

The Center for Interest Measurements Research

الذى اسس فى عام ١٩٦٢ فى جامعة مينسوتا بالولايات المتحدة • ويكفى أن نشير الى أن واحدا من هذه المشروعات يتضمن دراسة تمتد ٢٥ عاما لأكثر من الف تلميذ عمرهم ١٥ سنة • ومن المتوقع أن تزيد هذه الدراسات من قيمة الاختبار • وقد نقل هذا المقياس الى اللغة العربية عطية محمود هنا واستخدم فى عدد من البحوث المصرية والعربية •

ثانيا : اختبار كيودر للميول المهنية :

Kuder Preference Record-Vocational

يعتبر هذا الاختبار أكثر مجموعة اختبارات كيودر للميول شهرة أو استعمالا • والواقع أن الصور المختلفة ، والطبعات المعادة والمعدلة لاختبارات كيودر للميول تمثل عائلة من الأدوات التى تتخذ لقياس الميول سبلا وزوايا مختلفة ، كما أنها مصممة لأغراض مختلفة ، الا أن أهمها هو هذا الذى نتكلم عنه • وقد وضع هذا الاختبار كيودر ، واتبع منهجا مختلفا ، بل مضادا ، لمنهج سترونج فى اختبار وتصحيح عناصر الاختبار • لقد كان هدفه الرئيسى هو بيان الميل النسبى فى عدد صغير من المجالات الواسعة ، وليس فى مهن بذاتها ، وقد صيغت عباراته أصلا وجمعت مبدئيا على أساس صدق المحتوى • وقد اتبع هذا تحليل شامل لعناصر الاختبار من تطبيقها على جماعات من تلاميذ الثانوى والكبار • وكان الهدف من تحليل العناصر هذا هو انشاء مجموعات عناصر ذات اتساق داخلى عال وارتباط منخفض بالمجموعات الأخرى وقد تحقق هذا الهدف بشكل جيد فى معظم مقاييس الاختبار •

وعناصر اختبار كيودر من نوع الاختيار الاجبارى الثلاثى ، وقد تعرفنا عليه فى بعض اختبارات الشخصية من قبل _ وعلى المفحوص أن يختار بين الأنشطة الثلاثة المذكورة ذلك النشاط الذى يحبه اكثر كما يبين النشاط الذى يلقى عنده اقل تفضيل ، ومن امثلة هذه الثلاثيات ما يلى ، وهى مختارة من

الصورة العربية للاختبار التى أعدها الدكتور احمد زكى صالح (ص ٧ من كراسة الأسئلة) ٠

١٦ ـ تعمل نموذجا من الصلصال ٠
١٧ ـ تكتب مقالا عن طريقة اقناع الناس وتوجيههم ٠
١٨ ـ تعمل ملقنا فى تمثيلية يقوم بها هواه ٠
١٩ ـ تحترف الطب ٠
٢٠ ـ تحترف النحت ٠
٢١ ـ تحترف الصحافة ٠

يعطى الاختبار عشرة مقاييس ميـــل بالاضافة الى مقياس تحقـــق Verification scale للكشف عن الاهمال ، أو سوء الفهم ، واختيار الاجابات المقبولة والمرغوب فيها اجتماعيا ولكنها غير متوقعة ٠ وتتضمن مقاييس الميول العشرة الميول الآتية : الخلوى ، والميكانيكى ، والحسابى ، والعلمى ، والاقناعى ، والفنى ، والأدبى والموسيقى ، والعدالة الاجتماعية، والكتابى ، وقد وضع كيودر للاختبار معايير خاصة بكل من الذكور والاناث لتلاميذ الثانوى ، والجامعة ، والكبار ٠ وتخطط الدرجات الكلية على مجالات الميول العشرة على قائمة معايير مئينية ٠

وتدور معاملات ثبات اختبار كيودر حول ٩٠ر٠ ، كما يبدو الثبـات باعادة الاختبار بعد عام أو اقل مرضيا ، أما بعد فترات أطول فليس هناك قدر كاف من المعلومات عنها ، وما هو موجود يشير الى أنه ، وخاصـة حالة تلاميذ الثانوى ، كانت هناك تغيرات فى مجـالات الميول المرتفعة والمنخفضة عندما كانت الفترة بين الاختبارات تمتـد الى عدد من السنين (١٨ : ٤٧٥) ويبدو أن كلا من اختبار ستـرونج وكيودر عرضة للكذب أو التحـريف ، ويزيد هذا بالنسبة الى اختبار كيودر نظرا لطبيعة عنـاصره الاكثر وضوحا وشـفافية ٠

اما دراسات صدق الاختبار فقد أجريت أساسا بمقارنة الاختبار بمحك الرضا فى العمل ٠ ومن أكبر هذه الدراسات الطولية الشاملة تلك التى أجريت على ١١٦٤ تلميذا أعطوا اختبار كيودر فى الثانوى ، ثم أعطوا استبيانا

للرضا فى العمل بعد عشر سنوات ، وعند ذلك كان عدد ٧٢٨ من هـــؤلاء ملتحقين باعمال « متسقة » مع انماط ميولهم الأصلية ، و ٤٣٦ فى اعمـــال « غير متسقة » مع هذه الأنماط • وكانت نسبة العاملين الراضين فى اعمالهم فى الجماعة المتسقة هى ٦٢٪ ، بينما كانت النسبة هى ٣٤٪ فى الجماعة غير المتسقة ، ومن ناحية اخرى كانت نسبة العاملين غير الراضين ٨٪ فى الجماعة المتسقة ، و ٢٥٪ فى الجماعة غير المتسقة (١٨) • وقد ايدت هذه النتائج دراسات اخرى فى ذات الاتجـــاه •

وقد تطور حديثا عن اختبار كيودر اختبار آخر هو اختبار كيـودر لمسح الميل العـــام (٤٠) • Kuder General Interest Survey

وهو عبارة عن مراجعة وامتداد للاختبار الأصلى فى الاتجاه الأدنى من الميول وقد صمم للاعمار من نحو ١٢ حتى ٢٤ ، ويستعمل لغة ابسط والفاظا اسهل من الاختبار السابق ، ولهذا فهو لا يتطلب مستوى رفيعا من القراءة • وقد وضعت معايير مئينية امريكية لأكثر من عشرة آلاف ولد وبنت ، ولعينات قليلة من الكبار وما زال البحث جاريا للتحقق مما اذا كانت هذه الأداة تناسب الكبار كما تناسب من هم أصغر سنا ، وفى هذه الحالة قد ينتهى الأمر الى أن يحل اختبار كيودر لمسح الميل العام محل الاختبار السابق • وتشير الدراسات التى أجريت على الاختبارين أن الارتباطات بينهما تعادل معاملات الثبـات الخاصة بهما ، اى انها تشـــير الى درجة التشابه الكبير فى قيـــاس الميول عنـد الكبـــار •

وثمة صورة اخرى لاختيار كيودر للميول المهنية وهى (٤١) :
Kuder Occupational Intrest Survey (KOIS)

اختبار كيودر لمسح الميل المهنى ، وقد تطور هذا الاختبار وانشىء على أساس المحك وبطريقة متشابهة لتلك التى اتبعها سترونج فى اختباره ، كما سبق بيانه ، الا انه فى هذه الصورة لم يستعمل جماعة الرجال بصفة عامة كما هو الحال عند سترونج ، وبدلا منها فان درجة المفحوص على كل مقياس مهنة يعبر عنها ارتباط بين نمط ميوله ونمط ميول جماعة مهنية معينة • ويمكن الحصول حاليا على درجات لعدد ٧٩ مهنة و ٢٠ مجالا من مجالات الدراسة

الجـامعية للذكور ٠ و ٦٥ مهنة و ٢٥ مجـال دراسة جامعية للاناث ٠ وتتنوع المهن التى يتناولها هذا الاختبار تنوعا كبيرا فى مستواها فهى تتراوح بين الخباز وسائق الشاحنة (سيارة النقل) من طرف ، والكيماوى والمحامى من الطرف الآخر ، وقد ساعد حذف جماعة الرجال أو النساء بصفة عامة ، أو جماعة المرجع ، على اتساع المجال الذى تستوعبه هذه الأداة الواحدة ٠ ويتضمن هذا الاختبار بالاضافة الى مقاييس المهنة ومقاييس مجالات الدراسة الجامعية مقياس تحقق وثمانية مقاييس تجريبية تستعمل الآن لأغراض البحث العلمى فقط ٠

وكما هو فى اختبارات كيودر كلها فان عناصر الاختبـار مصاغة فى صورة الاختيار الاجبارى ، ويغطى محتوى العناصر مجالات الميول المتضمنة اختبار كيودر للميول المهنية مضافا اليها مجالات أخرى مثل تفضيل العمل مستقلا ، وتجنب الصراع ، الميل للمواقف المألوفة والمستقرة ٠

وقد أثبت كيودر بتحليل احصائى موسع لثلاثة آلاف مفحوص أن نظام المقاييس المهنية المشتقة من جماعات المرجع الخاصة بالرجال بصفة عامة ٠ كما تظهر الدراسات التى أجريت على ثبات هذا الاختبار وصدقه أنها مرضية الى حد بعيد ، وقد اعتمدت هذه الدراسات على ٣٠ مقياسا مهنيا ، ولا يزال أمام الاختبار ومقاييسه المهنية الأخرى مجـال لدراسة صدقها وخاصـة بالنسبة الى محك خارجى ، (١٨ ، ٤١) ، ولن تستقر قيمة هذا الاختبار النهائية الا بعد أن تتأكد بالبحث وبالتطبيق فى مجالات عملية متنوعة ٠

أما الصورة العربية لاختبار كيودر الميول المهنية ، والتى سبقت الاشارة اليها ، فقد قام باعدادها وصياغتها أحمد زكى صالح ، وتتكون من كراسة للاختبار نفسه ، وورقة يضع المفحوص عـلامة × فى الخانة الأكثر تفضيلا وأخرى فى الخـانة الخاصة باقل تفضيل لأساليب النشاط الثلاثة ، ثم تحسب الدرجات الخام الخاصة بكل ميل بعد أن يتأكد المصحح من صدق الاجابة ، ثم تحول الدرجات الخام الى مئينيات ويرسم تخطيط الميول على بطاقة تخطيط خاصة بالبنين وأخرى خاصة بالبنات ٠

وللاختبار كراسة تعليمات تتضمن معلومات عن وصف الاختبار وهدفه

والميول التي يقيسها ، ومفاتيح التصحيح ومقياس الصدق ومفتاحه (ص) ، ثم وصف لاستخراج معايير الاختبار ، وبيانات عن ثباته وصدقه . وقد حظى هذا الاختبار باهتمام علمي واسع في مصر ، واستعمل ولا يزال في عدد من البحوث في مجال الميول المهنية (١ : ٢٦١ ـ ٢١٤) .

ثالثــا : اختبــار لي ـ ثــورب The Lee — Thorpe Inventory

بينما يمثل اختبار سترونج الأساس التجريبي في بنائه وتصحيحه ، ويمثل كيودر التنوع والتعدد ، نجد أن اختيار لي ـ ثورب يمثل اتجاها ثالثا في بناء اختبارات الميول . هذا الاختبار يتكون من مجموعة من الأمثلة التي اختيرت ونظمت على أساس الحكم وليس على الأساس الاحصائي . هـذا الاتجاه الذي يمكن أن يسمى الاتجاه المنطقي ، مضاد لاتجاه سترونج الذي يعتمد على النتائج الاحصائية . ويشبه هذا الاتجاه المنطقي الأسـلوب الذي يتبع في بناء اختبارات القدرة والاستعداد حيث تعين مجموعة كبيرة جدا من المواقف ثم تختار عنصر عشوائيا من هذه المجموعة الضخمة .

وقد بدا لي ثورب بوصف المهن الموجودة في قاموس العناوين المهنية Dictionary of Occupational Titles الذي يصف ويصنف كل المهن والأعمال تقريبا في المجتمع الأمريكي ، واختيرت في ستة مجالات مهـام تمثل ثلاثة مستويات للمسئولية : مرتفع ، ومتوسط ، منخفض وتذكر أوصـاف موجزة للاعمال في أزواج ، ويبين المفحوص ما يفضله من واحد من زوج ، وتمثل الدرجة التكرار النسبي للاختبارات في أقسام : الشخصية الاجتماعية، والطبيعية ، والميكانيكية ، والأعمال ، والفنون ، والعلوم .

ويختلف هذا الاختبار عن اختبار سترونج في أن الاختبار الأصلي وتصنيف عناصره يعتمد كلية على أوصاف للمهنة وليس على دليل واقعي على أن أشخاصا في المهنة يحبون نشاطا معينا كما يختلف عن اختبار كيودر في أن التجميع الأول للعناصر بني على أساس تحليل منطقي وليس على أساس العزل الاحصـائي للعوامل وعلى الرغم من أن التصنيف الأصلي للعناصر اعتمد كليا على أساس منطقي فقد أجريت دراسات ارتباطية تالية بهدف رفع

مستوى التجانس ، وعند مراجعة الاختيار استبعدت العناصر التى لاتشترك مع بقية العناصر فى القسم الواحد (٢٧) ٠

الميول والشخصية : اتجاهات وتوقعات :

كانت الدراسات الأولى فى الميول تعتمد اعتمادا كبيرا على المقارنة الواقعية الخالصة لجماعات مهنية على أساس عناصر مختارة بطريقة عشوائية غالبا ، أى أنه لم يكن هناك نظرية خاصة بطبيعة الميول ، ولا بأنماط الميول التى يجب أن نوليها عناية خاصة ، وكذلك لم تكن هناك نظرية عن تكوين المهن والأعمال ، وكانت تفسيرات درجات اختبارات الميول تتم على أساس عملى غير نفسى (٢٧ : ٤٣٦) ٠ وقد تغير الاتجاه فى قياس الميول منذ بداية نشر سترونج اختباره ، ويمكن القول ان اختبار سترونج واختبار كيودر أديا الى بداية نظرية فى الميول ٠ وقد وجد جيلفورد وزملاؤه (٣١) أكثر من عشرين عامل ميل يبدو أن معظمها يعكس أساليب شخصية عامة وليس اتجاهات مهنية ٠ وأضاف جيلفورد الى ما هو معروف من ميول مثل الميكانيكية والعلمية وغيرها ، أضاف المغامرة فى مقابل الأمن ، والتذوق الجمالى ، والمسايرة الثقافية أو النظام ، والعدوان وأبعاد ميول أخرى كثيرة وكذلك أدت الدراسات الطولية ودراسات الحالة الى اعطاء صورة أكثر وضوحا عن اهمية الميول ومغزاها ٠ أى أن هناك تطورا ونموا متوقعا فى اتجاه نظرية متكاملة عن الميول ٠

وثمة اتجاه آخر نام أيضا هو التفسير المنظم للميول فى ضوء مفاهيم أساسية للشخصية ٠ فبينما أتى على الميول حين من الدهر كان ينظر اليها فيه على أنها عمليات اشتراط تقوم على الصدفة ٠ نجد أن الميول الآن ينظر اليها على أنها تعبير عن حاجات عميقة وعن أنماط توافق معينة ٠ وربما تكون اختبارات الميول أكثر كفاءة من عدد كبير من الأساليب الأخرى فى قياس الشخصية نظرا لتنوع محتواها وزيادة تقبل الشخص لها ٠

وقد كشف عدد كبير من الدراسات عن ارتباطات ذات دلالة بين الميول المهنية المقاسة وجوانب أخرى للشخصية ٠ فمن هذه الدراسات ما أظهر مثلا أنه فى حالات ادمان الخمر ، والجنسية المثلية ، والعصاب ، وغيرها من

نواحى العجز الذهانية كانت تسود كل جماعة من هذه الجماعات انماط ميول مهنية خاصة مميزة لها (٢٥ ، ٢٨) . وبينت دراسات اخرى اجريت على تلاميذ المدارس والجامعات والعمال الصناعيين والمرضى الراشدين ان الاستهداف للحوادث يرتبط ارتباطا ذا دلالة بمقاييس معينة على اختبار سترونج ، وكذلك ثبت أن هناك ارتباطا بين درجات اختبار سترونج وكيودر واختبارات الشخصية مثل اختبار مينسوتا المتعدد الأوجه وغيره (١٨ : ٤٩١)

ومن جهة أخرى فانه من المعترف به الآن أن اختيار مهنة غالبا ما يعكس حاجات الشخص العاطفية الأساسية ، وأن التوافق المهنى يمثل جانبا أساسيا فى توافق الشخص العام فى الحياة . ذلك لأن الشخص عندما يختار مهنة ، فانه الى حد بعيد ، يختار أساليب التوافق وأنماط الحياة والأدوار التى يحس انها قريبة اليه ومناسبة له . وعلى هذا فان قياس الميول المهنية ، وبصفة خاصة تحديد الجماعات المهنية التى يشارك الفرد أعضاءها ميولهم واتجاهاتهم ، يصبح امرا هاما فى فهم مختلف الشخصيات . ومما يؤيد التأكيد الحالى لاتخاذ الاختبارات المهنية مؤشرات للشخصية ذلك الاختبار الذى وضعه هولاند (٣٠) Holland Vocational Preference Inventory وقد استعمل فيه العناوين المهنية لقياس خصائص الشخصية . وما على المفحوص فى هذا الاختبار الا ان يوضح ما اذا كان يحب أو يكره كل واحدة من ثلاثمائة مهنة . وعلى الرغم من أن هذا الاختبار قد اعتمد اعتمادا كبيرا على صدق المحتوى الداخلى ، فان مقاييسه قد اظهرت اختلافات دالة بين عينتين متناظرتين احدهما تتكون من افراد مصابين بأمراض نفسية ، وتتكون الأخرى من أفراد عاديين ويعطى الاختبار عشرة مقاييس للشخصية ، منها النشاط البدنى والنزعة العقلية ، والمسئولية ، والمسايرة وما زالت قيمة هذا الاختبار رهن البحث والاستعمال العملى الا أنه يمثل اتجاها جديدا فى دراسة الشخصية عن طريق الميول .

من هذا يتضح انه من المتوقع أن تتطور نظرية خاصة بالميل ، وأن يمثل قياس الميول مجالا حيا فى تطور الاختبارات وفى البحث العلمى ، وأن تنمو وتتنوع الدراسات حول العلاقة بين الميول والشخصية ، وان يتجه تفسير اختبارات الميول الى مزيد من العمق والدقة .

مراجع الباب السادس

١ ـ احمد زكى صالح : علم النفس التربوى · القاهرة : مكتبة النهضة المصرية (ط ١٠) ، ١٩٧٢ ·

٢ ـ احمد زكى صالح : الأسس النفسية للتعليم الثانوى · القاهرة : مكتبة النهضة المصرية (ط ٢) ، ١٩٧٢ ·

٣ ـ احمد عبد العزيز سلامة : تطبيق اختبار تفهم الموضوع على حالات مصرية · رسالة ماجستير ، كلية التربية جامعة عين شمس ، ١٩٦٢ ·

٤ ـ سيد احمد عثمان : المسئولية الاجتماعية · القاهرة · مكتبة الأنجلو المصرية ، ١٩٧٣ ·

٥ ـ سيد أحمد عثمان : علم النفس الاجتماعى التربوى · القاهرة : مكتبة الأنجلو المصرية ، ١٩٧٠ ·

٦ ـ سيد أحمد عثمان ، فؤاد أبو حطب : التفكير ، دراسات نفسية ، القاهرة · مكتبة الأنجلو المصرية ، (ط ٢) ١٩٧٧ ·

٧ ـ سيد محمد غنيم : مدى صلاحية اختبار بقع الحبر لرورشاخ لقياس الذكاء · رسالة ماجستير ، كلية التربية جامعة عين شمس ، ١٩٥٥ ·

٨ ـ سيد محمد غنيم ، هدى عبد الحميد برادة : التشخيص النفسى : دراسات فى اختبار الرورشاخ ، القاهرة : دار النهضة العربية (حـ١)، ١٩٦٥ ، (حـ ٢) ١٩٦٦ ·

٩ ـ سيد محمد غنيم ، هدى عبد الحميد برادة : الاختبارات الإسقاطية · القاهرة ، دار النهضة العربية ، ١٩٦٤ ·

١٠ ـ فاروق أبراهيم أبو عوف : دراسة مقارنة لمعرفة مدى الافادة من اختبار رورشاخ (بقع الحبر) للتمييز بين حالات العصاب القهرى

والأسوياء · رسالة ماجستير ، كلية التربية جـــامعة عين شمس ، ١٩٧٢ ·

١١ ـ فاروق ابراهيم أبو عوف : مدى صلاحية اختبار رورشـــاخ لقياس الابتكارية · رسالة دكتوراه ، كليـــة التربية جامعة عين شمس ، ١٩٧٦ ·

١٢ ـ فاروق ابراهيم أبو عوف : دراسة مقارنة لبعض سمات الشخصية بين القادة وغير القادة باستخدام اختبـــار بقع الحبر لرورشاخ · الكتاب السنوى فى علم النفس ، الجمعية المصرية للدراسات النفسية، المجلد الرابع ، القاهرة · مكتبة الأنجلو المصرية ، ١٩٨٥ ·

١٣ ـ فؤاد أبو حطب : التحليل العلمى للسلوك الخلقى · الكتاب السنوى فى التربية وعلم النفس (تحرير سعيد اسماعيل على) · المجلد الأول ، القاهرة : عالم الكتب ، ١٩٧٣ ·

١٤ ـ فؤاد أبو حطب : التفضـــيل الفنى وســـمات الشخصية · المجـلة الاجتماعية القومية ، يناير ١٩٧٣ ·

١٥ ـ لويس كامل مليكه ، محمد عماد الدين اسماعيل : الشخصيةوقياسها · مكتبة النهضة المصرية ، ١٩٥٩ ·

١٦ ـ نجيب اسكندر ابراهيم ، لويس كامل مليكه ، رشدى فام منصور : الدراسة العلمية للسلوك الاجتماعى · القاهرة : مؤسسة المطبوعات الحديثة ، ١٩٦٦ ·

١٧ ـ يوسف الشيخ ، جابر عبد الحميد جابر : سيكولوجية الفروق الفردية القاهرة : دار النهضة العربية ، ١٩٦٤ ·

18. Anastasi, A. Psychological Testing. New York, Macmillan, 1982.

19. Atkinson, J.W. (ed.) Motivation, fantasy action, and society. Princeton, N.J. Van Nostland, 1958.
20. Atkinson, J.W. and Feather, N.T. A theory of achievement Motivation. New York : Wiley, 1966.
21. Barkes, R.G., Kounin, G. and Wright, H.F. (eds.) Child behaviour and development. New York : McGraw-Hill, 1943.
22. Beck, S.L. Rorchach's test. I : Basic processes, New York : Grune and Stratton, 1944.
23. Beck, S.L. Rorchach's test. III : Advances in interpretation. New York : Grune and Stratton, 1952.
24. Bell, H.M. The Adjustment Inventory. Palto Alto, Calif : Consulting Psychologists, 1934-1939.
25. Carnes, G.D. Vocational interest characteristic of abnormal personalities. J. of Counsel, Psychol. 1964, II, 272-279.
26. Combs. A.W., a Snygg, D. Individual Behaviour, New York : Harper, 1959.
27. Cronbach, G.L. Essentials of Psychological Testing, New York : Harper, 1984.
28. Drasgow, J. and Carkhuff, R.R. Kuder neuro-psychiatric Keys before and after psychotherapy. J. of Counsel. Psychol. 1964, 11, 67-71.
29. Flanagan, G.C. Factor analysis in the study of personality. Stanford : Stanford Univ. Press, 1935.
30. Goldner, R.H. Individual differences in whole-part approach and flexibility-rigidity in problem solving. Psycholo. Monogr. 1957, 71, No. 21.
31. Guilford, L.P. et al., A factor analysis study of human interests. Psychol. Morograph, 1954, 68, No. 41.
32. Hartshorne, H. and May, Studies in deceit. New York : Macmillan, 1928.
33. Hartshorne, H. and May, M. Studies in service and self-control, New York : Macmillan, 1929.
34. Hartshorne, H. and May, M. Studies in the organization of character, New York : Macmillan, 1930.

35. Hathaway, S.R., and Mckinley, G.C. Minnesota Multiphasic Personality Inventory : Manual for administration and scoring. New York : Psychological Corporation, 1967.

36. Holland, J.L. A personality Inventory employing occupational titles. J. of Appl. Psychol., 1958, 42, 336-342.

37. Holtzmann, W.H., et al. Inkblot perception and personality-Holtzmann Inkblot Technique. Austin, Texas : Univ. of Texas Press, 1961.

38. Klein, G.S. and Schlesinger, H.L. Perceptual attitudes toward instability : I. Prediction of apparent movement experiences from Research response. J. Personol., 1951, 19, 289-302.

39. Krech, D., Crutchfield, R.S. and Ballachey, E. L. Individual In Society. New York : McGraw-Hill, 1962.

40. Kuder, G.F. General Interest Survey, Manual Chicago : Science Research Associates, 1964.

41. Kuder, G.F. Kuder Occupational Interest Survey : General Manual. Chicago : Science Research Associates, 1966.

42. Kunce, J.T. Vocational Interests and accident proneness. J. of Appl. Psychol., 1967, 51, 223.

43. Lindzey, G. and Borgatta, E.F. Sociometric measurement. In Lindzey, G. (ed.) Handbook of Social Psychology. Cambridge, Mass. Addison-Wesley, 1954, Vol. I, Ch. II.

44. Luchins, A.S. Mechanization in Problem solving. Psychol. Monogr., 1942, 54, No. 6.

45. Masling, J. The influence of situational and interpersonal variables in projective testing. Psychol. Bulletin, 1960, 56, 65-85.

46. Masling, J. Differential introduction of examiners and Rorcharch responses. J. of Consult. Psychol., 1965, 29, 198-201.

47. McClelland, D.C., et al. The achievement motive. New York: Appleton-Century, 1953.

48. McClelland, D.C. The achieving society. New York : Free Press, 1961.

49. Moreno, G.L. Who Shall Survive ? Washington : Nervous and Mental Diseases Publishing Co., 1934.

50. Murray, H.A., et al. Explorations in personality. New York: Oxford Univ. Press, 1938.

51. Myers, I.B. The Myers-Briggs Type Indicator : Manual Princeton, N.J. : Educational Testing Service, 1962.

52. Peters, H.N. and Campell, G.T. Diagnosis of training needs of B-29 mechanics from supervisory ratings and self-report. AFPTRC Tech. Memorandum 55-12, 1955.

53. Roe, A. Two Rorchach Scoring Techniques : The inception Technique and the basic Rorchach. J. Abrom. Soc. Psychol., 1952, 47, 263-264.

54. Rogers, C.R. Counseling and Psychotherapy, Cambridge, Mass; Riverside, 1942.

55. Rogers, C.R. Client-Centered Therapy. Cambridge, Mass. : Riverside, 1951.

56. Ruggles, R. and Allport, G.W. Recent Applicat 1 of the A.-S. Reaction Study. J. Abnorm. Soc. Psychol., 19 , 34, 518-528.

57. Sarason, S.B. The clinical interaction with special reference to the Rorchach. New York : Harper, 1954.

58. Strong, E.K. and Campbell, D.P. Manual for strong Vocational Interest Blank Stanford, Calif : Stanford Univ. Press, 1966.

59. Taylor, Janet A.A. personnality scale of manifest Anxiety. J. Abnorm. Soc. Psychol., 1953, 48, 285-290.

60. Wallen, R.W. Clinical Psychology. New York : McGraw-Hill, 1956.

خاتمـــة

تقويم أدوات التقويم

عرضنا فى فصول هذا الكتاب المفاهيم الأساسية التى تســتخدم فى الحكم على صلاحية الاختبارات وادوات التقويم بوجه عام ، ثم طبقنا هذه المفاهيم على امثلة رئيسية من اختبارات الذكاء والقــدرات والاستعدادات والتحصيل والشخصية والميول ٠

ونعرض فى هذه الخاتمة نموذجا يصلح لتقويم الاختبــارات وادوات التقويم الأخرى يتألف من ٢٨ بندا ، وفيه تتطلب البنود رقم ١ الى رقم ٩ وصفا بسيطا للاختبار يمكن أن نحصل عليه عادة من كراسة تعليماته والدليل الخاص به ، ويتطلب البند رقم ٦ وصف الطابع العام للاختبار ٠ كما أن البند رقم ٧ (تاريخ النشر) ليس على درجة كبيرة من الأهمية لأن بعض الاختبارات القديمة قد تكون ممتازة ٠ وأى اختبار قديم يجب أن يفحص بعناية ودقة ، فقد تكون بعض اسئلته غير ملائمة لتغير العصر ، وقد تكون معاييره غير مفيدة ، وقد يكون دليله غير كامل ٠ وفى هذا البند يجب أن نسجل أى طبعات جديدة أو أى تعديلات طرات على الاختبارات ٠

ويحتوى النموذج المشار اليه على البنـد رقم ١١ الذى يتطلب وصف الاختبار من حيث عناصر وأسئلته وتصحيحه من اجل تكوين انطبــاع عام عن الاختبار ولا شك أن هذا البند هو الذى يحدد مدى جاذبية الاختبــار للمفحوص وللمرجل العادى ٠ اما البند ١٢ يتطلب تحديد مقاصد المؤلف سواء كانت انتقاء او توجيها أو علاجا او تقويما مدرسيا ، وبالتالى يحدد بهدف المؤلف الى قياسه بالاختبار ٠ وقد تتضمن بعض كراسات التعليمات الأسلوب الاحصائى الذى استخدمه المؤلف فى انتقاء الأسئلة أو تحليلها وحساب صدق تلك الأسئلة ٠

ويتضمن البند ١٣ حكما على تعليمات الاختبار فى ضوء الأسس العامة التى عرضناها فى الفصل الثانى من الكتاب ، اما البند ١٤ فيتطلب أن نقوم بتحليل نظرى لأسئلة الاختبار للحكم على ما تتضمنه من قدرات أو خبرات او عادات أو سمات شخصية ، وكيف تؤثر هذه العوامل كلها أو بعضها فى

درجة الاختبار . وهذا التحليل ضرورى لكل درجة من درجات الاختبارات الفرعية التى يحتويها الاختبار وبالطبع نحن لا نستطيع أن نحـــدد جميع العوامل المسهمة فى درجة الاختبار . ولكن مثل هذا العمل التحليلى يثير عددا من الأسئلة يمكن أن تستخدم فى تفســـير نتائج الاختبار . ويجب أن يسجل التقرير ما تدل عليه الدرجة فى الاختبار سواء كان حكمنا هذا يتفق مع مقصد المؤلف أو لا يتفق ، كما يوضح العوامل غير المرتبطة التى تؤثر فى الدرجة وقد تفسدها . ويمكن للباحث أن يلجأ الى تطبيق الاختبار على نفسه أو على مفحوص آخر لديه خبرة بالقياس النفسى ثم يلاحظ الباحث أداءه فى الاختبار ، أما فى حالة اختبار المفحوص لنفسه فان الملاحظة تصبح من نوع التأمل الباطنى introspection ، ورغم الاعتراضات المنهجية على هذه الطريقة الا أنها تفيد كثيرا فى اعطـــاء كثير من الدلالات عن طبيعة الاختبـــار .

وتدل البنود من ١٦ ـ ١٩ ما يتوفر عن صدق الاختبـــار من بيانات تجريبية ، وذلك للحكم على مدى صلاحية هذه البيانات لأغراض الاختبار ، أما البنود من ٢٠ ـ ٢٢ فتشمل بيانات ثبات الاختبار ، ويتصل البند ٢٣ بثبات الاختبار من حيث فاحصه ومصححيه ، والمعايير من حيث دقة تفسير الأداء فى الاختبار كما يدل عليها البندان ٢٤ ، ٢٥ .

ويشير البند ٢٦ الى تلخيص العروض الناقدة للاختبار فى المؤلفات والمجلات المتخصصة . أما البند ٢٧ وموضوعه « تقويم عـــام للاختبار » فيتضمن تلخيصا عاما ونهائيا لنواحى الضعف والقوة مع اعتبار النواحى العلمية والفنية جميعا ، ومن الملائم مقارنة الاختبار بغيره من الاختبارات التى تقيس نفس الوظائف العامة ، بالاضافة الى تسجيل الملاحظات العامة للفاحصين والمفحوصين وتعليقاتهم الخاصة عليه .

وهكذا يتوافر للفاحص قائمة تحليلية للاختبـــار بحيث يســـتطيع الفاحصون ـ وخاصة أولئك المشاركون فى البرامج الشاملة للتقويم سواء التربية أو الصناعة أو القوات المسلحة ـ ان يختاروا الاختبـــار المناسب للغرض المناسب فى الوقت المناسب ، وبأقل فاقد ممكن فى الجهـد والوقت والمال .

نموذج لتقويم الاختبارات

١ ـ عنوان الاختبار ٠

٢ ـ اسم مؤلف الاختبار ومترجمه (أن وجد) ٠

٣ ـ الناشر ٠

٤ ـ صور الاختبار والجماعات التى يمكن أن تطبيق عليها ٠

٥ ـ المعالم العملية (اية خصائص عملية تتصل بالأداء فيه أو تصحيحه)٠

٦ ـ نوع الاختبار (اختبار قدرات أو شخصية مثلا) ٠

٧ ـ تاريخ النشر ٠

٨ ـ التكاليف ، كراسة التعليمات والأسئلة ، ورق الاجابة ٠

٩ ـ الزمن المطلوب ٠

١٠ ـ الغرض الذى من أجله نحكم على الاختبار ٠

١١ ـ وصف الاختبار من حيث عناصره ، وتصحيحه ٠

١٢ ـ غرض المؤلف واساس اختبار عناصر الاختبار عنده ٠

١٣ ـ دقة التعليمات ووضوحها ، والتدريب المطلوب لمن يقوم بعملية اعطاء الاختبار ، ودرجة تقنينها ٠

١٤ ـ الوظائف أو السمات العقلية الممثلة فى كل درجة ٠

١٥ ـ تعليقات وملاحظات عن تصميم الاختبار وبنائه (ورقة الأسئلة وطريقة الاجابة) ٠

١٦ ـ الصدق التنبؤى (المحك ، وعدد ونوع الحالات التى استخدمت فى حسابه ، والنتيجة) ٠

١٧ ـ الصدق التلازمى (المحك ، وعدد ونوع الحالات المستخدمة ، والنتيجة) ٠

١٨ ـ اية ادلة تجريبية أخرى تبين ما يقيسه الاختبار (صدق التكوين الفرضى) ٠

١٩ ـ تعليقات وملاحظات عن صدق الاختبار بالنسبة لأغراض معينة سيستعمل من اجلها الاختبار ٠

٢٠ ـ ثبات الاختبار بحساب معامل الاستقرار (الطريقة التى اتبعت للوصول اليه وعدد ونوع الحالات ، والنتيجة) ٠

٢١ ـ ثبات الاختبار بحساب معامل التكافؤ (الطريقة التى اتبعت للوصول اليه ، وعدد نوع الحالات ، والنتيجة) ٠

٢٢ ـ ثبات الاختبار بحساب معاملات الاتساق الداخلى أو اتساق الأسئلة (الطريقة التى اتبعت للوصول اليها ، وعدد ونوع الحالات ، والنتيجة) ٠

٢٣ ـ موضوعية الاختبار من حيث التصحيح تفسير الدرجات ٠

٢٤ ـ المعايير (نوعها ، وعدد ونوع الحالات) ٠

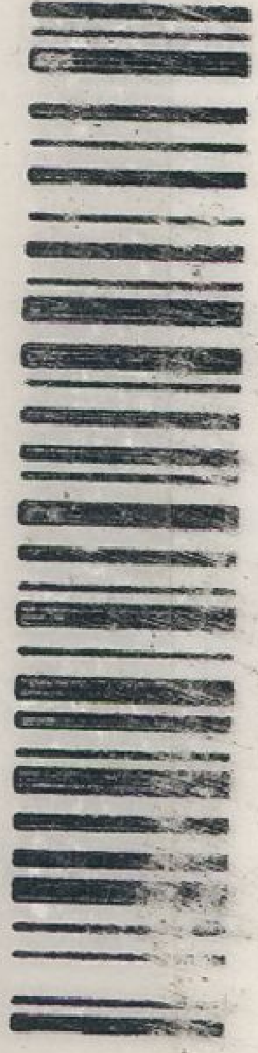
٢٥ ـ تعليقات وملاحظات عن دقة فئات الاختبار وصحة المعايير للغرض الذى سيستعمل الاختبار من أجله ٠

٢٦ ـ تعليقات وملاحظات من قاموا باستعراض للاختبار وتقويمه ونقده ٠

٢٧ ـ تقويم عام للاختبار ٠

٢٨ ـ المراجع ٠




Bibliotheca Alexandrina
06558938